

aer

**THE ALBERTA JOURNAL OF
EDUCATIONAL RESEARCH**

VOLUME XXX

NUMBER 4

DECEMBER 1984

**PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON**

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Published in March, June, September, and December by the
Faculty of Education, The University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

EDITOR: Walter H. Worth

CONSULTING EDITORS

D. C. Bjorkquist
University of Minnesota

P. E. Bryant
University of Oxford

B. M. Chung
Han Rim University,
Republic of Korea

A. Côté
University of Ottawa

D. Fitzgerald
The University of New England,
Australia

D. B. Harrison
The University of Calgary

N. L. Hersom
The University of Saskatchewan

M. Laferrière
McGill University

T. Morrison
University of Manitoba

J. Richardson
University of New Brunswick

P. Robinson
Atlantic Institute of Education

T. Rogers
University of British Columbia

R. E. Traub
Ontario Institute for
Studies in Education

K. Watson
University of Reading

R. J. Wilson
Queen's University

The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Alberta Education Index*, *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, *ERIC/Current Index to Journals in Education*, *Multidisciplinary Index to the World's Social Sciences Literature*, and *Social Science Citation Index*; appropriate articles are abstracted in *Cognitive Development Abstracts*, *Educational Administration Abstracts*, *Educational Technology Abstracts*, *Language Behavior Abstracts*, *Multicultural Education Abstracts*, *Psychological Abstracts*, *Research into Higher Education Abstracts*, *School Organization and Management Abstracts*, and *Sociology of Education Abstracts*, and reviewed by the *Journal for Research in Mathematics Education*.

The subscription rate is \$18.00 per year for individuals, \$24.00 per year for institutions. Add \$6.00 for delivery outside of Canada. Single copies are \$6.00 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, The University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.



The Alberta Journal of Educational Research

Volume XXX, Number 4

December, 1984

CONTENTS

Teachers' Perspectives on Inservice Education	239
<i>M. Pansegrau</i>	
One Last Word List	259
<i>L. Gunderson</i>	
Communication Skills in Immersion Programs	270
<i>S. T. Carey and J. Cummins</i>	
Student Teachers' Perceptions of Their Competence Versus the Perceptions of University Supervisors and Sponsor Teachers	284
<i>C. F. Staab</i>	
An Evaluation of the Diagnostic Efficiency of the Wechsler Intelligence Scale for Children-Revised	299
<i>H. H. Mueller, G. J. Mancini and R. H. Short</i>	
Computer-Assisted Instruction in College Composition	311
<i>P. Evans, N. Mickelson and N. C. Smith</i>	
Delegation: The Neglected Aspect of Participation in Decision Making	320
<i>W. K. Hoy and D. A. Sousa</i>	

PERSPECTIVES

The Theory-Practice Relationship in Teacher Education: Need There Be Any?	332
<i>P. J. Miller and G. D. Taylor</i>	
TABLE OF CONTENTS, VOLUME XXX	340

FACULTY OF EDUCATION
The University of Alberta

MORAG PANSEGRAU

University of Alberta

Teachers' Perspectives on Inservice Education

The purpose of this study was to determine teachers' perspectives on inservice education. It was postulated that the beliefs and activities—the who, what, when, where and why—of those who participate in, or are expected to participate in, inservice education would fall into patterns and that the portrayal of such patterns would provide some insight into the meaning held by teachers of inservice education.

Data were collected over a twelve-month period using the field research strategies of semi-structured, in-depth interviews and participant observation. The findings of the study reveal that teachers classify inservice education into the following four discrete groups of learning activities, each of which they perceive differently—Group One: mandatory, formal inservice education activities offered in conjunction with a compulsory program change; Group Two: other types of mandatory, formal inservice education activities; Group Three: formal inservice education activities of a voluntary nature; and, Group Four: non-formal inservice education activities. Based on the findings of the study, recommendations for school systems and universities are given.

Introduction and Need for the Study

It has long been recognized in English-speaking countries that preservice teacher education represents but the minimum prerequisite for entry into the teaching profession. Teachers must continue to develop professionally throughout their teaching career. Consequently, inservice education activities have been offered since the late 19th century (Richey, 1957; Stephens & Hartman, 1978; Bacon, 1980). These activities are seen as a means of ensuring that practicing teachers are exposed to new techniques, programs and other innovations and are given assistance with classroom-related and subject-related areas of concern.

The current knowledge explosion and rapid rate of technological growth, the aging and relatively static teacher population, the current public dissatisfaction

Morag Pansegrau holds a PhD in Educational Administration and, at the time this article was submitted, was Coordinator of the Committee for Improvement of Teaching and Learning, University of Alberta. She is now Senior Instructor in Business Education, Alberta Vocational Centre, Edmonton, Alberta.

with the products of the educational system, the ever-expanding role being assigned by society to educational institutions and the rapidly changing lifestyle of many of the segments of our society, all contribute to the necessity of ensuring that teachers are involved in an ongoing program of professional development. As Bacon (1980) points out:

even the most excellent preservice program is only a preparation for entry into the profession, and cannot last a whole career because the context teachers work in changes. (p. 1)

Indeed, according to Houle (1980), a zest for knowledge and lifelong learning are key traits that qualify an individual for membership in a profession. Many professions have come to the realization that their members must individually and collectively accept the obligation to continue to learn and that programs to achieve this purpose must be established (Harrington, 1979; Frandson, 1980; Houle, 1980).

Unfortunately, as much of the literature over the past two decades acknowledges, inservice education has been relatively ineffective as a vehicle for introducing educational reform into classrooms. Despite vast expenditures by government, by school boards, by professional councils, by universities and by teachers (Fennell, Hill, & Thiessen, 1980; *ATA News*, 1981), large numbers of teachers appear to view such activities with indifference or hostility, attend reluctantly, if at all, and do not appear to change their classroom behavior in the ways anticipated (Fullan & Pomfret, 1977; Cooper & Hunt, 1978; Brayne, 1980). Consequently, "many of our grand plans for educational change have failed to have major impact or have been co-opted" (West, 1975, p. 35). As Howey and Joyce (1978) point out, inservice education does not appear to be in the best of health, despite Bruce's (1980, p. 140) prediction that, "in the 1980s innovation and growth will be found in the inservice education of practising teachers." Indeed, it has been described as the "slum of American education" (Wood & Thompson, 1980, p. 374). A pertinent question is posed by King, Hayes, and Newman (1977):

Why do teachers, who almost universally appear eager to improve their professional performance, frequently respond with disdain or outright hostility to local efforts to "inservice" them? (p. 686)

In order to combat teachers' inferred indifference to their professional development as indicated by their reluctance to attend inservice education activities organized for them by others, and their unwillingness when they do attend to adopt the information presented, a variety of solutions has appeared in the literature. The solutions currently in vogue include involving teachers in the planning stages, conducting needs assessments and manipulating various externals such as the timing, format and location of activities. What is commonly assumed by those writing about and organizing inservice education is that teachers who attend inservice activities will transfer the learning to the classroom and it will be observed in the changed routine of the classroom. When teachers do not immediately adopt the new behaviors, it is further assumed that the delivery format needs to be altered. The key to successful inservice education thus is seen as the discovery of new and motivating models for the delivery of inservice education, and the literature is replete with criteria, guidelines, models and lists of shortcomings and recommendations, all of which, it is inferred, will result in more effective inservice education.

It does not appear to have been seriously considered that teachers may attend inservice education activities for reasons other than to obtain information which they will use to bring about an immediate change in their classroom behavior, or

that teachers may deliberately choose not to adopt the new behavior. What have not been determined are teachers' perspectives on inservice education.

Perhaps, before suggesting a plethora of antidotes to combat teachers' alleged indifference or hostility towards inservice education, teachers' perspectives on inservice education should be determined. Although there exists a vast array of articles and texts on the topic of inservice education, the field lacks a formal, systematic research base and thus "until inservice teacher education is examined more closely, we can only talk about what we think we know" (Cruickshank, Lorish, & Thompson, 1979, p. 31). In particular, inservice education research has given little attention to descriptive studies (Cruickshank, Lorish, & Thompson, 1979, p. 31). Thus, this study was designed as an inductive, descriptive and analytical study which would permit the researcher to determine how teachers perceive the world of inservice education but would avoid imposing upon teachers the preconceptions of others in relation to teacher inservice education.

Consequently, in this study effort was devoted to an understanding of the total phenomenon of inservice education by determining the beliefs and activities—the who, what, when, where and why—of those who participate in, or are expected to participate in, inservice education activities. It was posited that their beliefs and activities would fall into patterns and that the portrayal of these patterns would provide an insight into the meaning for teachers of inservice education.

Methodology and Procedures

Traditionally, research in the field of education has adopted research methods associated with the natural or physical sciences. As Blumer (1969) states:

the predominant procedure is to take for granted one's premises about the nature of the empirical world and not to examine these premises. . . . Instead of going to the empirical social world in the first and last instances, resort is made instead to a priori theoretical schemes, to sets of unverified concepts, and to canonized protocols of research procedure. (pp. 32-33)

In recent years, however, there has been growing disenchantment with the processes and results of traditional educational research methods and an increasing interest in what, according to Spradley and McCurdy (1972) and Bogdan and Biklen (1982), are generally labeled ethnographic methods. These methods, as West (1977, p. 61) states, "allow one to understand how conceptions held by people shape their behavior, at least in part, and how such conceptions and behavior change over time."

In an exploratory study of a relatively unresearched topic, a methodology and procedures which would permit the researcher to determine what Bruyn (1966) terms "inner perspective," or the understanding of people from their own frame of reference, were sought. The symbolic interactionist perspective (Blumer, 1969, 1978), a methodological stance that respects the nature of the empirical world, guided the research; and, an inductive approach in which the researcher enters the field with an area of interest, some general questions to be answered, but no more, was adopted.

Teachers' perspectives on inservice education were gathered through the field research strategies of in-depth interviewing and participant observation. The major data gathering strategy was a series of arranged in-depth elite and specialized interviews (Dexter, 1970) which used an open-ended, semi-structured approach in order to focus the interview yet permit the interviewee to define the situation, structure his account of it and introduce his notions of what he regarded as rele-

vant. Longhand and shorthand notes were used to record the views of participants and the observations of the researcher.

The researcher had access to teachers and inservice education activities in two school systems. Over a twelve-month period, 77 arranged in-depth interviews were conducted with 86 teachers in 11 schools. In addition, four days of teacher conventions and two professional development days organized for the teachers included in the study were attended. Informal conversations with many teachers and those responsible for organizing inservice education activities for teachers were a part of the data-gathering.

The general approach to inductive research as outlined by Glaser and Strauss (1965, 1966, 1967) and Glaser (1978) prevailed. This permits the researcher to ground the research and theory in the reality he/she is studying and allows "important concepts, basic categories and significant hypotheses" (Glaser & Strauss, 1966, p. 57) to emerge. The data collection process is guided by the principle of theoretical sampling, theoretical saturation and the constant comparative method. The researcher collects, codes and analyzes the data and decides what data next to collect and where to find it; evidence is collected from other comparative groups to check out whether the initial evidence is correct. Theoretical saturation is the criterion for judging when to stop sampling the different groups pertinent to a category; that is, no additional data are being found whereby the properties of the category can be developed. Thus data collection, coding and analysis is ongoing, points the direction for future sampling and determines the focus of future interview probes.

The field was entered with a minimum of predetermined assumptions and a general question to be answered, namely, "What are teachers' perspectives on inservice education?" Although the data collection, coding and analysis steps outlined in this study and often found in the literature may appear to have a linear quality, in reality they often occur simultaneously. Entry may have been under negotiation in one school, interviews underway in another, while data collected from another school were under analysis. The data obtained from the first interviews revealed certain views held by teachers about inservice education and suggested tentative hypotheses to be explored. The first group of interviews was conducted with high school teachers. After the responses were coded and analyzed it was decided to interview teachers at other grade levels to compare their views. The views of teachers in promoted positions were sought to determine if they varied from those of classroom teachers. Interviews were directed to teachers highly receptive to organized inservice activities as well as to those with less enthusiasm for such activities. Data were collected from teachers in elementary, junior high and senior high schools, and from administrators. Both male and female teachers with a variety of subject area specialties and at various stages in their careers were included in the sample. The research also sought to determine if those teachers choosing not to volunteer to participate in the study were radically different from those volunteering; it was judged that there was no discernible difference.

The data collected were analyzed using the coding and category development tactics suggested by Turner (1981) and taking into consideration the suggestions of Glaser and Strauss (1967), Glaser (1978) and Bodgan and Biklen (1982). The four elements of a perspective as outlined by Becker, Geer, Hughes and Strauss (1961, p. 436) guided the organization of the categories and the presentation of patterns. The data were grouped, finally, into 12 categories, and from these categories emerged patterns which reveal teachers' shared perceptions of inservice education (see appendix).

Summary of the Findings

It was the intent of this study to determine teachers' beliefs and activities in relation to inservice education in order to ascertain if there were common patterns which would permit a greater understanding of the phenomenon of teacher inservice education. It was posited that those writing about and organizing teacher inservice education and those attending inservice education activities may be operating from different premises. An analysis of the data collected indicates that teachers share certain perceptions relating to inservice education. These shared perceptions reveal a discrepancy between how teachers perceive inservice education and how others assume teachers to perceive inservice education, assumptions which usually form the rationale for the various inservice education activities organized for teachers.

VALUE OF AN ACTIVITY	essential				
	rewards must equal sacrifices				
	worthless				
GROWTH (higher level needs) ↑ NEEDS SATISFACTION (reasons for attending various activities) ↓ DEFICIENCY (lower level needs)	to obtain information that will be used to effect a major change in class-room behavior				
	to enjoy a mentally stimulating break from routine				
	to obtain recognition that a worthwhile job is being performed				
	to associate with adults instead of children and exchange information with colleagues				
	to become acquainted with the latest developments in the field of education.				
	to obtain confirmation that present practices are appropriate and acceptable				
	to obtain information and materials relating to present teaching assignment which can be immediately adopted (requires no major change in ideology or current practices)				
	non-attendance as attending will not satisfy a need				
		group one: mandatory, formal inservice education activities offered in conjunction with a required program change	group two: other types of mandatory, formal inservice education activities	group three: formal inservice education activities of a voluntary nature	group four: non-formal inservice education activities
		NOMOTHETIC < ----- ORIENTATION ----- > IDIOGRAPHIC			

Figure 1. Groups of inservice education activities: A model of inservice education as perceived by teachers

The findings of the study reveal that there is not one single perspective on inservice education held by teachers but rather that teachers perceive inservice education to fall into four discrete groups of learning activities, each of which they perceive differently. Their decision whether or not to participate, their reasons for participating and their assessment of the value of an activity vary according to the particular group of activities in question.

Teachers' shared perceptions of inservice education are portrayed in Figure 1, "A Model of Inservice Education as Perceived by Teachers."

Teachers classify inservice education by types of activities which range from formal activities organized for teachers by others with a nomothetic or institutional orientation, to activities of a non-formal, ongoing nature, that tend to be individualized and teacher initiated, and have an idiographic orientation. Teachers participate in activities of both a formal and non-formal nature in order to satisfy certain needs which range from lower level deficiency needs to higher level growth needs. Although participation in an activity may satisfy more than one need, usually there is a dominant need which the teacher desires to have satisfied and which accounts for his attending one activity rather than another. The value of a particular activity is determined by the individual, his needs at that moment in time, and whether attending the activity satisfied those needs. In general, the activities are regarded as essential, worthless, or worthwhile if the rewards derived from attending equal or exceed the sacrifices or costs that are incurred by attending. Following is a synopsis, by groups of activities, of teachers' perspectives on inservice education, as illustrated in Figure 1.

Group One: Mandatory, Formal Inservice Education Activities Offered in Conjunction with a Compulsory Program Change

These activities are usually held during the regular school day and are attended by those who are presently teaching the subject or who will be responsible for this area in the future. Teachers attend formal inservice education activities organized in conjunction with a compulsory program change in order to obtain information which they will use to effect a major change in their classroom behavior (whether such a change actually occurs is outside the scope of this study). They consider these activities to be essential if they are to interpret and implement new programs.

If you are going to use a new curriculum, inservice is an issue, a big deal. Teachers know they are going to have to teach the curriculum and they had better go to the inservice.

In social studies the written prepared units will do the job of telling the teacher what to teach . . . but the one-day inservice will bring it all together. You need both parts to be effective.

They prefer sessions held during the school day (with students released from classes) and they want relevant information and materials presented by classroom teachers with expertise in the area.

I feel much better this year with a day off for social studies inservice, you're not looking at the clock after 4 p.m. . . . It's the first time for a one-day inservice in school time, in depth, done by a teacher not a publisher, and it makes it worthwhile.

It's of benefit having no students on the day of the social studies inservice as then you don't have to prepare for subs, worry about the class clown, tie up all the loose ends on your return to school.

Group Two: Other Types of Mandatory, Formal Inservice Education Activities

Included in this category are formal inservice education activities held during the regular school day which teachers are obligated to attend (students are released from classes). Such activities include professional development days, teachers' conventions and sessions organized by the school or school system.

Much of the criticism leveled by teachers against this type of activity seems to have its roots in the fact that there is no long-term plan for continuing teacher professional development. Rather, activities are offered in a haphazard, piecemeal fashion, usually the topics are of a "one of" nature, tend to be repetitious, are offered at only one level and do not cater to the different needs of teachers at various stages in their career.

No attempt is made to match the needs of teachers with what is organized, no attempt at all.

P.d. is offered piecemeal. You are made to go to activities but there's no plan when you compare it to other professions . . . they insist on certain things and perhaps they make people more aware of what they should be looking at (re p.d.). We don't do that in Education yet.

Teacher inservice education needs follow an identifiable pattern which changes over time—a pattern that is consistent with the literature and research on adult learning and development, the nature of professional learning and the concerns of teachers.

There is a pattern, a change. At first you are insecure, you have to do a great job in many areas and you attend inservice randomly, with not much selection. As a beginning teacher I would go to just anything . . . then after I had taught a few years I became more selective.

It appears that at the beginning of their teaching career, teachers feel insecure and overwhelmed and desire formal inservice education activities that will help them cope with their teaching assignments.

When I first started I felt incomplete and I went to everything. Now I feel more competent and I know how to teach reading, I have used four different language arts series . . . so I don't go to language arts sessions.

As they become more experienced, they gain in confidence and knowledge and often prefer more thought-provoking presentations.

I think the beginning teacher tends to look for more things that could help him in the classroom be a better teacher as he is groping. A more experienced teacher feels more confident, goes to more abstract, inspirational things, probably that's why I prefer conventions now.

However, whenever they make a subject or grade change or are promoted, initially they function as neophytes in the new position.

After the first degree, the first few things you go to you want someone to tell you what to do on Monday morning. Once you've got the day-to-day under control you are more theory oriented. . . . Same applies when you make a change.

The teachers included in this study believe that there are many ways in which teachers can continue their professional development and that attending formal inservice activities is but one way.

Over ten years I have seen every approach from the evangelical to the practical, hands-on and now it seems that there is a need on p.d. days to let me come to school in my jeans and have time to myself in my personal school environment.

Teachers say, let me as a professional say where I want to put the time. . . . I really do p.d. activities at other times, not necessarily on official professional development days.

Although attending these activities permits teachers to satisfy certain needs, the dominant reason for attending is not a desire to obtain information which they will use to effect a major change in their classroom behavior.

I don't necessarily go and say, "Hey, I am going to get new ideas for my classroom."

I think it is more of a recharging experience and if you believe what you hear you may change your behavior but I don't think, honestly, that happens.

Consequently, these activities may be considered to be less than successful when judged by the criterion of an overt change in the classroom behavior of teachers. On the other hand, the activities may be considered eminently successful when judged by the various criteria identified by the teachers included in this study. Although to organizers who expect to see an overt behavior change result from teacher participation, the needs identified by teachers may appear to be inconsequential, almost frivolous, and the benefits derived nebulous, to teachers these needs are very important and their satisfaction exceedingly desirable, if not essential, for their mental health and sense of well-being. Indeed, although teachers would like to see changes made, many consider these activities to be worthwhile, regardless of topic or presenter.

You get a great deal of value from organized inservice activities, even if you go to bad ones . . . they're stimulating, even despite the speaker.

Following are the various needs identified by teachers that are satisfied by attending mandatory, formal inservice education activities that do not accompany a compulsory program change. The various needs range from lower level needs of a safety, security and belonging nature to those at a higher level that promote growth.

A. *Non-attendance.* Some teachers choose not to attend these activities, even though they are held during school hours and students are released from classes. These teachers consider school days without students to be essential, but do not approve of the ways in which these days currently are organized. As adults and professionals, they want the freedom to choose if, and when, to participate in formal inservice education activities. They believe that on days when students are released from classes teachers should have the right to engage in whatever types of learning activities they, not others, consider will be most beneficial to them and to their students.

We never say: What would you like to do to be a better teacher?

There should be a choice of activities—do what you want whether it is socializing or workshops.

Teachers go to conferences, workshops, we think are important.

B. *To obtain information and materials relating to present teaching assignment which can be immediately adopted (requires no major change in ideology or current practices).* Teachers are constantly seeking new ideas, techniques, behaviors and materials which they can use to enliven their present classroom practices but which require no fundamental change in ideology or current practices. At these sessions they desire practical, classroom-related information which they can put into practice immediately and which requires only those materials currently available in their schools. They prefer presentations given by individuals who are, or were until recently, classroom teachers and who understand the realities of life in the classroom. Too, they want hands-on activities which permit them to become involved and handouts, as such take-away information greatly enhances the likelihood of their using the information presented.

I prefer highly technical workshops, with teaching techniques, lesson plans—I want concrete activities for the classroom.

I want hands-on inservice using materials that are available in our school. I don't want to be given ideas that you can't carry out yourself.

I want hands-on, small group activities given by presenters who have dealt with the same age level and situation as you . . . you forget what the classroom is like . . . something to pick up and take with you—the only way it will come back to the classroom.

C. To obtain confirmation that present classroom practices are appropriate and acceptable. It has been documented that teachers accept the findings of educational research that are neutral or confirm present practices, otherwise findings tend to be ignored (Clifford, 1973, p. 27). The same appears to be true for formal inservice educational presentations.

We believe what is relevant and what backs up our position; we don't remember or distort what is not acceptable to us.

Teachers want to attend positive presentations which enhance their self-concept by reassuring them of the appropriateness of their present practices.

You listen to someone and they tell you to do this and that and you are doing it and you say, "Hey, that's what I'm doing," and that's good. There's a lack of pats on the back for teachers, there's no recognition.

I felt so good when the presenter mentioned certain things I was doing.

They indicated that unless they were in a crisis situation which necessitated an immediate change in their classroom behavior (for example, a compulsory program change), they will not adopt a new approach, even if they consider it to have merit.

Unless it's an area I don't know too much about I go for confirmation, not a conscious desire to change unless I feel shaky in the area. I wouldn't adopt right away unless I felt dissatisfied with my approach, I've not got the time.

D. To become acquainted with the latest developments in the field of education. Teachers wish to keep current, and desire presentations which will expose them to new knowledge. Although they admit that they probably will not adopt the new concept, they state that they are open to the seeds of a new idea being sown. However, they want the new knowledge to germinate and come to fruition if and when they decide that the concept is acceptable and the timing appropriate. If they change their behavior, usually it is after several exposures over an extended period of time.

To change behavior takes years, but inservice acquaints teachers, stimulates them and if they choose to follow up, that's their decision.

I get ideas from speakers, and encouragement, but I think it is a slow assimilation.

It may even take more than one inservice on a topic before you consider using it.

Indeed, even if they desire to adopt a new idea, they would wait until "next time around"; there are simply too many daily tasks that must be performed.

Even if you change because of inservice, you probably will not change that year or the next; you might adopt a couple of ideas a couple of years later . . . changing a program is really hard. I try to, but it puts more pressure on you. . . . It's odd, whether you are a first year or a ten year teacher, none of us like the pressure and the pressure does not lessen.

I tend to not change behavior, partly because of lack of time . . . I would adopt at the beginning of the next time I was teaching the unit. It's bad enough just getting through the day-to-day with 101 things to do . . . I've not got two weeks to set up a new program.

It was also suggested that teaching does not attract innovators.

Part of the reason people get into teaching is that they are very conservative, rational people—if they were radical, they would not last, or like, teaching. . . . You tend to go to inservice on, for example, dyslexia, listen, return to the classroom and pick up at the chalkboard where you left off, and make no attempt to use the information on learning/reading disabilities.

Some pointed out that it is easier to bring about change if others in the school are knowledgeable and supportive.

It's difficult to apply new ideas, concepts, if other staff members are not interested.

It was suggested that changing behavior after an inservice which accompanies a compulsory program change is easier because the teacher does not have to convince others of the need to change.

To adopt, you must justify new ideas to bosses, colleagues, incorporate them into your present teaching methods. It's different with social studies inservice as the curriculum is in use and you will be evaluated on it.

However, as several teachers mentioned, those involved must feel the need to change.

There must be some sort of dissatisfaction with what is, and if you go to a workshop and you feel dissatisfied, then you are tempted to change . . . unless you can create that need for change, it is very hard. The problem of inservice is that if you have to inservice people, it must be linked to a need, and you have to create that need for change and if it is not created, the whole inservice is a waste of time.

We wanted to change a couple of teachers in a certain school, make them more empathetic. We gave a session on discipline. The ones already empathetic became more so, felt positive, expanded, but the ones we wanted to change, it was like water off a duck's back—that was us imposing our view of their needs on them and they did not perceive it as a need and did not respond to the session.

In the final analysis:

Whether you use, implement the idea depends on what it is, how it's presented and if you were interested in the first place.

E. To associate with adults instead of children and exchange information with colleagues. The isolation of the classroom teacher appears to be a cause of great concern to teachers. Attending organized inservice education activities ensures that teachers obtain a much needed mental health break from the regimented world of the school and the isolation of the classroom.

I like kids . . . but I need to escape to an adult level intellectually.

I think that getting out of the classroom and exchanging ideas is a good 50 percent of the value (of formal inservice education activities) and value enough.

Consequently, they like activities to be arranged in a manner that permits and encourages social contacts.

I see inservice activities as the perfect opportunity to talk with other teachers in the same field as you are. I see that as a significant benefit.

Too, attending activities reminds them that they belong to a profession.

Despite the fact that often the offerings are not as useful as one would hope, I think there is an aura of fraternity, of fellowship, that surfaces at this type of gathering which makes it worthwhile, regardless of what happens . . . you go about your working day and you forget that you belong to a profession.

F. To obtain recognition that a worthwhile job is being performed. Teachers suffer from a lack of positive feedback concerning the importance and effectiveness of their performance. Consequently, they seek presentations which give them a

mental pat on the back, supply the desired recognition and encourage them to return to the classroom and persevere with their teaching assignment.

It's confidence building in some ways. It's hard to get recognition (in teaching), it's one of the problems of the profession. The kids don't say it, adults don't see you perform and you just don't know . . . inservice serves that type of purpose, it's confidence building.

To a great extent we are appreciated by the community but are not told about it, we just hear the negative and it's nice to be told. Keynote speakers tell us that . . . and we take the pats on the back and say maybe we're not doing such a bad job.

G. *To enjoy a mentally stimulating break from routine.* It is difficult for teachers to escape mentally from the confines of the classroom, and inspiring, thought-provoking presentations which permit teachers to extend their mental horizons beyond the four walls of their own classroom and subject specialization are welcomed. Such presentations help them to keep their own classroom and its concerns in perspective and they return rejuvenated and encouraged.

Inservice on any topic is refreshing, a change of pace, a little spark.

Going to meetings, workshops can open up new vistas.

From a shot-in-the-arm session you get a sparkle and the kids benefit.

Group Three: Formal Inservice Education Activities of a Voluntary Nature

These activities usually are organized by the school, the school system, the specialist councils of the teachers' professional association or the university, and may be held during school hours, in the evening, on the weekend or during school holidays. For this group of activities, teachers do not believe that benefits will accrue regardless of topic or presenter, as they do for mandatory formal activities that do not accompany a compulsory program change. Although teachers may attend these activities for any of the reasons listed in Figure 1, in general they seek information and materials which will help them with their present teaching assignment. They prefer to have these activities scheduled within normal school hours since they do not feel they have surplus time or energy for activities held out of working hours.

The criterion they use to determine if they should attend these activities is "will the rewards equal the sacrifices?" If the activity is held during the school day, the sacrifices or costs incurred by attending are the loss of teaching time and the work of preparing for a substitute teacher. If the activity attended does not provide them with the anticipated rewards, they suffer feelings of guilt because they consider that they have unnecessarily deprived their students of valuable teaching time.

You should leave inservice feeling excited and wanting to go back to the classroom and try it out yet the last inservice here I left and felt badly about the waste of time. I would rather have been teaching.

If the activity is held outside of normal school hours, attending means forgoing part of their leisure time and they want reassurance that the benefits derived will compensate this loss.

Depends on the value of the session. If I can be guaranteed it will be worthwhile, then I don't mind evening or weekend.

That teachers believe they do not have the time or energy to attend activities held out of school hours agrees with McClusky's (1973) concept of margin. McClusky (1973) considers that a necessary condition for learning is the availability of margin, or surplus power or energy left over after a person has handled his load. Given society's never-ending expectations of the educational system, the knowledge explo-

sion and the increasing number of teachers who are in the middle years, it is not surprising that many teachers believe that by the end of the school day, week or year, they do not have any surplus energy for formal inservice education activities.

I don't mind devoting some of my own time but when you ask me to go at four o'clock, I don't have much to put into it . . . teachers are worn out by the end of the school day.

As one teacher explained:

Teachers fear they (inservice activities) are a waste of time because they frequently have been. They want a guarantee that when they spend time it will pay off . . . also there's the problem of conservation of resources, a personal thing: How much can I spend and still survive as a human being? They do have the time and energy if they can be guaranteed a payoff—no time means no time to waste on inservice if they are guaranteed only a 20 per cent chance of it being worthwhile.

Group Four: Non-Formal Inservice Education Activities

Non-formal inservice education activities are ongoing, often casual, involvements which can occur at any time and in any location. Usually they are teacher initiated and involve an individual teacher or a small group of colleagues. The most frequently cited non-formal activity is daily in-school contacts with colleagues, in particular with those colleagues who have the same subject specialization.

Yet it goes on all the time—when you think about it, professional development goes on all the time in the staffroom.

Peer-based consultation, one-to-one or small group, is more important than formal inservice workshops.

I pick up more in the staffroom during the course of the year than at workshops.

Other non-formal activities are reading professional literature, involvement with new programs, having student teachers and watching others teach, and participating in extra-curricular activities that have spin-offs to the classroom.

Reading books and periodicals is of great value. I've become very familiar with the ATA library . . . I try to visit there once every three months.

My professional development comes from working with colleagues in school and from my choice of extra-curricular activities. Reading too . . . a lot comes through the introduction of new programs.

You can get professional growth by watching other teachers in action in their schools. Also, working with student teachers is a professional activity and every teacher could benefit from it.

Teachers indicated that they regularly use non-formal inservice education activities to keep current, obtain additional classroom-related information and materials, and find solutions to classroom-related problems. Thus it appears that, in general, it is from their involvement in these activities rather than formal inservice education activities, that teachers obtain information and materials which they will use to effect major changes in their classroom behavior and so improve their effectiveness. Since participation in these activities directly affects their performance in the classroom, teachers consider these involvements to be, at a minimum, worthwhile if not essential. The importance to teachers of non-formal activities, in particular in-school contacts with colleagues who have the same areas of specialization, are revealed by the comments of teachers who are the only specialists in their subject area in the school.

What I do enjoy (at inservice activities) is meeting with other librarians in the district because the business of being a librarian is relatively lonely. There may be five social studies teachers in the school so that if you have a problem you can kick it around a bit, but when

you are a one-teacher specialist such as the librarian or the home economics teacher, you are alone.

Conclusions, Recommendations and Implications

Conclusions

Traditionally, it has been assumed that by participating in formal inservice education activities teachers will obtain information which they will use to bring about an immediate change in their classroom behavior and so maintain or improve their effectiveness. According to the findings of this study, such an assumption is not always accurate. Teachers do endeavor to keep current and provide their students with the best learning experiences possible, but they use a variety of methods to attain this goal, and consider in-school contacts with colleagues, rather than involvement in formal inservice education activities, to play a vital role.

Thus, based on the findings of this study, it appears that formal inservice education activities which do not accompany a mandatory program change are not an appropriate method to use to effect a major change in the classroom behavior of teachers. These activities, however, do serve a useful purpose, as attending them permits teachers to satisfy certain basic and essential needs in the affective and lower level cognitive domains. Consequently, such activities should be considered worthwhile and should continue to be organized for teachers.

Formal inservice education activities can, however, make a more important contribution than simply permitting teachers to satisfy certain basic needs. But first, there must be a long-term plan for teacher professional development. This should take into consideration the principles of adult learning, the developmental stages of adults, the nature and pattern of teachers' professional needs, and should include all of the parties involved in teacher professional development—the individual, the employer, the professional association and the university. Formal inservice education activities would comprise but one component of the plan, and would be organized in a manner which acknowledged and provided for individual as well as subject area differences. It should ensure that teachers have choice, choice among activities and also choice as to whether or not to attend. It should involve teachers, not relegate them to passive roles, should recognize that they bring to sessions valuable knowledge and experience, and should permit them to enjoy and benefit from social contacts with colleagues.

It is interesting to note that school systems traditionally have put their money, time and energy into organizing formal inservice education activities and have often given little encouragement to, or ignored, teachers' involvement in non-formal activities. Yet given the findings of this study, it appears that it is through their involvement with non-formal, not formal, activities that teachers are obtaining knowledge which they will use to bring about changes in their classroom behavior; they attend formal activities for other, but to them very important, reasons. If teachers are to be the beneficiaries of inservice education, it could be argued that it is their perspective which should determine the nature of inservice education.

Recommendations and Implications

Recommendations for School Systems. The district or central office of the school system should continue to organize formal inservice education activities for teachers and also provide some funds for individual professional development.

Teachers should be given release time to attend activities held during school time.

- The district office of the school system should acquaint the public with the nature of, and need for, continuing teacher professional development. In particular, the public should be made aware of the fact that teachers are involved in ongoing continuing education experiences and that many of the activities which contribute to an improvement in teacher effectiveness are of a non-formal, often apparently social, nature. Too, that on days when students are released from classes, teachers can be engaged in meaningful learning experiences without attending group activities held in local schools.
- All formal inservice education activities should be held within the school day, week and year. The school system should organize such activities or should permit teachers to attend such activities organized by others. Part, if not all, of the cost should be borne by the school system.
- When introducing compulsory program changes, the school system and/or the department of education should ensure that such program changes are accompanied by formal inservice education activities held during the school day, preferably with students released from classes. The presentations should be given by classroom teachers with expertise in the area who can help teachers interpret and implement the new curriculum and who can suggest appropriate classroom materials.
- The formal inservice education activities organized for teachers should include the following types of presentations:
 - a. Practical, classroom-related types of presentations preferably given by those with recent classroom experience in the area. The ideas presented should be capable of being immediately adopted by the teacher on his return to the classroom and should require only those materials that are presently available in the schools. The preferred format for such activities is small group, hands-on activities which permit teachers to become involved and share ideas with other participants. Handouts are appreciated as they remind teachers of the information presented.
 - b. Information sessions that acquaint teachers with the latest developments in the field of education.
 - c. Positive, stimulating, thought-provoking presentations that enhance teachers' self-concept. Such presentations may relate directly to the field of education or may be of a more general nature.
- Since in any group of teachers there will be those who are novices in an area and those with a great deal of experience, subject matter presentations should be offered at a variety of levels to cater to the different needs of teachers at different stages in their careers.
- Since teachers have indicated their preference for presenters who are classroom teachers and pointed out the lack of qualified presenters, the school system should encourage and reward, in some way, those teachers who express an interest in giving occasional presentations and who are considered to have the necessary expertise. It is suggested that these rewards need not necessarily be of a financial nature but could take the form of release time, recognition, reimbursement for expenditures, and so forth.
- It is suggested that the school system should not rely too heavily on imported speakers who have no knowledge of the local scene but simply give a prepared presentation and leave.
- When organizing activities, the school system should take the time of year into consideration. Teachers have indicated that at certain times of the year they prefer more subject-oriented presentations whereas at other times of the year they desire thought-provoking, mentally stimulating presentations.
- Formal inservice education activities should be organized in a manner which deliberately encourages interactions of a social nature.
- The formal inservice education activities organized by the school system should follow and support a long-term plan for teacher professional growth. The weakness of a "one of" workshop is not so much the fact that it represents only one presentation on a topic, but rather

that the single presentation has nothing whatsoever to do with the previous or the following session.

- The importance of non-formal inservice education activities should be acknowledged by the school system and involvement in these types of activities should be encouraged.
- Teachers should have the freedom to choose if, how, and when to participate in formal and non-formal inservice education activities. In particular, on school days when students are released from classes, they should have the freedom to participate in whatever types of activities they consider will best benefit both themselves as professionals and their students.
- If the school system wishes to have teachers adopt an innovation, the following suggestions, based on the comments of the teachers included in the study, should be taken under consideration:
 - a. Teachers must be exposed to the innovation on several occasions over an extended period of time in order to permit them to become thoroughly acquainted with, and accept, the proposed innovation.
 - b. The school system must not only expose teachers to the innovation but also must take steps to encourage teachers and affected others to adopt the innovation. Teachers included in the study indicated that it is much easier to adopt an innovation if other members of the staff also are willing to change. The present common practice of exposing only one member of the school staff to an innovation and expecting that individual to persuade the rest of the staff to change, is unrealistic. As one teacher said, "when others on staff are unfamiliar with a new idea, it fizzles out."
 - c. The school system must realize that even when teachers wish to change their classroom behavior, usually they do not do so until "next time around," thus the timing of the exposure to the innovation is important.
- Since teachers have indicated that they have no time and little interest in being involved in planning formal inservice education activities, the present practice of offering concurrent sessions is an appropriate way of attempting to meet the needs of different groups of teachers. When organizing such activities, it should be remembered that teachers desire not only different topics but also different levels of subject matter presentations, and that teachers wish to have the choice of participating in formal or non-formal activities.
- Many of the teachers included in the study used terms which suggest that they perceive themselves as being isolated and powerless members of a large, impersonal organization. It is suggested that the school system should make every effort to ensure that teachers are, and perceive themselves to be, important, contributing members of the organization who have some control over the various aspects of their professional life.
- Since many teachers mentioned feelings of isolation and indicated that they considered in-school contacts with colleagues to be an important form of non-formal inservice education, it is suggested that the school system should deliberately foster such teacher-initiated consultations.
- A special effort should be made by the school system to cater to the needs of teachers who are the only subject specialists in their schools, as such teachers feel particularly isolated and neglected. They are deprived of ongoing, in-school contacts with colleagues in the same area of specialization and often their needs are overlooked at formal activities.
- School systems should consider introducing voluntary teacher-to-teacher networks similar to those described in *IMPACT II* (Mann, 1982). The approach to inservice education described in *IMPACT II* capitalizes on an important finding of this study, namely, that in-school consultation with colleagues contributes to the professional growth of teachers and the alleviation of their concerns.
- It is suggested that the school system could use the instrument developed by the researcher (Figure 1) to ascertain the needs that teachers desire to have satisfied by attending the various groups of inservice activities, and the value attributed to each group. Such information could help avoid a mismatch between the goals of organizers and participants.

- At present, formal inservice education activities do not appear to be evaluated, at least not in any strict sense of the term. If, in the future, an evaluation of these activities is considered to be desirable, it is suggested that since many of the benefits derived by participants are of an intangible nature, both traditional and non-traditional evaluation methods should be used.

Recommendations for Universities. Some teachers considered the university to play a vital role in their professional development; but many considered it to be a degree-granting institution, a “salary upgrading place,” which did not play a significant role. It is a place “for me to use in any way I want,” to which they turn when they want in-depth knowledge or a guest speaker. But, when they seek information which will help them improve their performance in the classroom, they turn elsewhere because, in their eyes, professors lack credibility—they are too far and too long removed from the world of the school and the classroom teacher.

Don't mention the university in regards to inservice as it's the ivory tower; it's years since they've been in the classroom. I'd go to university to be brought up-to-date on things in fantasyland.

The presentations are fairly useless if you go back (to school) with the intent of using these ideas on Monday morning.

The most frequently made recommendation was that professors in the faculty of education should have recent and regular exposure to the world of the classroom teacher.

My big gripe . . . is that many people in the faculty of education have not been in the classroom for 20 years . . . it's not the same just coming to observe a student teacher. So if you train teachers, after 5 years you should be back in the classroom for one day a week. You have to teach these people—be responsible for a class—to see what it's like now. Professors have a total misconception of what the classroom is like now.

Also, it was suggested that the university could play a more positive and powerful role in developing the preservice-inservice continuum.

At university (preservice), no one talked to us about inservice, we just saw notices posted. I didn't know if you were supposed to go or not. I went to a couple of Saturday morning sessions and found them quite valuable but wished my friends had been there as I felt quite inadequate with teachers of eight to ten years' experience. The university should stress attending inservice activities.

Following are recommendations for the university, based on the findings of this study which, it is suggested, would permit the university to play a more active role in the continuing education of teachers.

- The university should offer a wide variety of short-term credit and non-credit, off-campus and on-campus courses similar to those evaluated by Unruh (1981) which are organized in accordance with teachers' perspective on inservice education.
- The university should take steps to deliberately foster the development of a preservice-inservice continuum. In particular, preservice education students should be exposed to the concept of lifelong continuing teacher education and the various types of formal and non-formal ways in which teachers can continue their inservice education. Undergraduate education students should be encouraged to attend formal inservice education activities as part of their teacher education training; attending such activities should help to develop a feeling of professionalism and collegiality.
- As Kersh (1979, p. 45) states, “college faculty still act as if inservice is ‘doing something to somebody or running something for somebody.’” The university should encourage an on-going, two-way exchange of ideas between university faculty and teachers in the classroom.
- University professors in the faculty of education should be involved in a continuing professional education program, a major component of which would be regular, in-depth classroom experiences.

Towards the Generation of Substantive Theory

Glaser and Strauss (1965, 1967) advocate the discovery of theory from data systematically obtained from social research, which they call grounded theory, as opposed to theory generated from a priori assumptions. They consider that grounded substantive theory:

can give participants in a situation a broader guide to what they already tend to do, and perhaps help them to be more effective in doing it. (Glaser & Strauss, 1967, p. 248)

Based on the findings, following is an attempt to predict and explain teachers' attitudes and behavior in relation to inservice education.

- Teachers obtain information which they use to effect major changes in their classroom behavior and so improve their effectiveness as teachers from their involvement in a variety of non-formal inservice education activities, not from their involvement in formal inservice education activities, apart from those activities which accompany a compulsory program change.
- Most teachers are not dissatisfied with their present classroom practices and attend formal inservice education activities which do not accompany a compulsory program change in order to obtain confirmation that their present classroom practices are acceptable and appropriate, not to obtain information which they will use to effect a major change in their present classroom behavior.
- Teachers consider formal inservice education activities that do not accompany a compulsory program change to be information sessions not sessions designed to immediately change their classroom behavior.
- Teachers consider formal inservice education activities that do not accompany a compulsory program change to have little relationship to their day-by-day classroom activities. They attend formal inservice education activities that do not accompany a compulsory program change in order to satisfy lower level cognitive needs or needs in the affective domain.
- Teachers readily adopt new techniques and materials which require no major changes in their present classroom behavior or ideology. They adopt innovations which require a major change in their classroom behavior in crisis situations only, or when they are compelled to do so because of a compulsory program change.
- Teachers do not voluntarily adopt an innovation which requires a major change in their classroom behavior until they have been exposed to the innovation on several occasions over an extended period of time.
- Teachers use a cost-benefit formula to determine whether or not to attend voluntary formal inservice education activities.
- There is an identifiable pattern to the inservice education needs of teachers which changes over time as the individual moves along his career path from neophyte to experienced teacher. When a teacher changes grades, area of specialization or is promoted, however, initially in the new position he tends to function as a neophyte.
- Teachers who have satisfied most of their inservice education needs at all levels from deficiency to growth seek new challenges by changing grades, areas of specialization or seeking promoted positions.
- Teachers who are deprived of ongoing, in-school contacts with colleagues who have the same area of specialization suffer from extreme feelings of isolation and are deprived of an important source of information for improving their classroom effectiveness.
- One cannot truly understand the world of another unless one lives in that world. Consequently, in order for university professors to be able to offer classroom-related assistance to teachers, they must regularly live in the world of the classroom teacher.

Appendix

Final Categories Developed From the Data

The following four elements of a perspective, as identified by Becker, Geer, Hughes, and Strauss (1961, p. 436), form the conceptual framework used to portray teachers' perspectives on inservice education:

- a. a definition of the situation in which the actors are involved,
- b. a statement of the goals they are trying to achieve,
- c. a set of ideas specifying what kinds of activities are expedient and proper,
- d. and a set of activities or practices congruent with them.

The element is identified to the left of the category number:

- (a) 1. Teachers' definitions.
- (b) 2. Attitude of teachers towards their own professional development.
- (c,d) 3. Attitudes of teachers towards formal inservice education activities.
- (c,d) 4. Teachers' reasons for attending formal inservice education activities that are not a part of a compulsory program change.
- (c,d) 5. Teachers' views on topics offered at formal inservice education activities.
- (c,d) 6. Changes in attitude over time towards formal inservice education activities.
- (c,d) 7. Teachers' views on attendance and timing of formal inservice education activities.
- (c,d) 8. Teachers' views on presenters at formal inservice education activities.
- (c,d) 9. Teachers' criticisms of formal inservice education activities.
- (c,d) 10. Non-formal forms of inservice education.
- (c,d) 11. Teachers' comments relating to inservice education and the role of district office.
- (c,d) 12. Teachers' comments relating to inservice education and the role of the university.

References

- Alberta Teachers Association. (1981). Committee proposes solutions to inservice problems. *ATA News*, 15(10), 1.
- Bacon, R. (1980). Inservice in Britain: INSET or inservice education and training? *Elements*, 12(3).
- Becker, H. S., Geer, B., Hughes, E., & Strauss, A. (1961). *Boys in white*. Chicago: University of Chicago Press.
- Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. Englewood Cliffs, NJ: Prentice-Hall.
- Blumer, H. (1978). Society as symbolic interaction. In A. Wells (Ed.), *Contemporary sociological theories*. Santa Monica, CA: Goodyear.
- Bogdan, R. C., & Biklen, S. K. (1982). *Qualitative research in education: An introduction to theory and methods*. Boston: Allyn & Bacon.
- Brayne, R. (1980). The principal and teacher professional development—some strategies. *Studies in Educational Administration*, 21.
- Bruce, M. G. (1980). Inservice education grows massively in Western Europe. *Phi Delta Kappan*, 62(2), 140-141.
- Bruyn, S. T. (1966). *The human perspective in sociology*. Englewood Cliffs, NJ: Prentice-Hall.
- Cicourel, A. V., & Kitsuse, J. I. (1963). *The educational decision-makers*. Indianapolis: Bobbs-Merrill.
- Clifford, G. J. (1973). A history of the impact of research on teaching. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally.

- Cooper, J. O., & Hunt, K. P. (1978). A cooperative approach to inservice training. *Viewpoints in Teaching and Learning*, 54(4), 61-69.
- Cruikshank, D., Lorish, C., & Thompson, L. (1979). What we think we know about in-service education. *Journal of Teacher Education*, 30(1), 27-31.
- Denzin, N. K. (1970). *The research act*. Chicago: Aldine.
- Dexter, L. A. (1970). *Elite and specialized interviewing*. Evanston: Northwestern University Press.
- Douglas, J. D. (1976). *Investigative social research*. Beverly Hills: Sage.
- Fennell, B. H., Hill, W. L., & Thiessen, S. J. (1980). *Teacher inservice training costs: A staff study*. Alberta: Alberta Education.
- Frandsen, P. E. (1980). Continuing education for the professions. In E. J. Boone, R. W. Shearon, & E. E. White (Eds.), *Serving personal and community needs through education*. San Francisco: Jossey-Bass.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research*, 47(1), 335-397.
- Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.
- Glaser, B. G., & Strauss, A. L. (1965). Discovery of substantive theory: A basic strategy underlying qualitative research. *American Behavioral Scientist*, 8, 5-12.
- Glaser, B. G., & Strauss, A. L. (1966). The purpose and credibility of qualitative research. *Nursing Research*, 15(1), 56-61.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago: Aldine.
- Goodman, P. S. (1977). *New perspectives on organizational effectiveness*. San Francisco: Jossey-Bass.
- Guba, E. G. (1981). *Criteria for assessing the trustworthiness of naturalistic inquiries*. Paper presented at the University Council for Educational Administration Graduate Student Seminar, University of Alberta, Edmonton, May 3-6, 1981.
- Haller, E. J. (1979). Questionnaires and the dissertation in educational administration. *Educational Administration Quarterly*, 15(1), 47-66.
- Harrington, F. H. (1979). *The future of adult education*. San Francisco: Jossey-Bass.
- Houle, C. O. (1963). *The inquiring mind*. Madison: University of Wisconsin Press.
- Houle, C. O. (1980). *Continuing learning in the professions*. San Francisco: Jossey-Bass.
- Howey, K., & Joyce, B. (1978). A data base for future directions in inservice education. *Theory Into Practice*, 17(3), 206-211.
- Kersh, B. Y. (1979). Needed: New ways of thinking and acting in staff development. *Journal of Teacher Education*, 30(1), 44-46.
- King, J. C., Hayes, P. C., & Newman, I. (1977). Some requirements for successful inservice education. *Phi Delta Kappan*, 58(9), 686-687.
- Mann, D. (1982). The impact of IMPACT II. *Phi Delta Kappan*, 63(9), 612-614.
- McClusky, H. (1973). An approach to a differential psychology of the adult potential. In M. Knowles, *The adult learner: A neglected species*. Houston: Gulf.
- Richey, H. G. (1957). Growth of the modern conception of in-service education. In N. B. Henry (Ed.), *In-service education for teachers, supervisors, and administrators*. Chicago: University of Chicago Press.
- Spradley, J. P., & McCurdy, D. W. (1972). *The cultural experience: Ethnography in a complex society*. Chicago: Science Research Associates.
- Stephens, T. M., & Hartman, A. C. (1978). Inservice education: Its murky past and uncertain future. *Viewpoints in Teaching and Learning*, 54(4), 1-3.

- Turner, B. A. (1981). Some practical aspects of qualitative data analysis: One way of organizing the cognitive processes associated with the generation of grounded theory. *Quality and Quantity*, 15, 225-247.
- Unruh, W. R. (1981). *An evaluation of the design, development and implementation of several course formats for teacher in-service in education*. Calgary: Faculty of Education, University of Calgary.
- Webb, S., & Webb, B. (1968). *Methods of social study*. New York: Kelley. (Original work published in 1932).
- West, W. G. (1975). Participant observation research on the social construction of everyday classroom order. *Interchange*, 6(4), 35-43.
- West, W. G. (1977). Participant observation in Canadian classrooms: The need, rationale, technique, and development implications. *Canadian Journal of Education*, 2(3), 55-74.
- Wood, F. H., & Thompson, S. R. (1980). Guidelines for better staff development. *Educational Leadership*, 37(5), 374-378.

LEE GUNDERSON

The University of British Columbia

One Last Word List

This article begins with a discussion of large corpus and pedagogical word lists. Pedagogical lists, generally developed from large corpus lists, are constructed to represent significant portions of the vocabulary in school texts. Teachers often teach the words on pedagogical lists so their students will know on sight most of the words they will encounter in written discourse. Over the last two decades many lists have been published as improvements to the "Dolch list of 220 words," the most widely used list, each claiming they represent more of the vocabulary in selected texts.

Two pedagogical lists are developed here, one based on the words produced by 80 second, fifth, eighth, and eleventh grade subjects writing three different compositions and the other produced by only the second grade subjects. The first list is as successful as well known published lists, including the Dolch, in accounting for substantial portions of the vocabulary of beginning reading texts. The second grade list, developed from a corpus of only 2,000 words, represents over half of the vocabulary in beginning reading texts.

The article concludes with the suggestion that researchers have done enough as far as pedagogical word lists are concerned; they are all equally successful in representing the vocabulary of beginning reading texts. The task remaining is to investigate the vocabulary of individual beginning reading texts, vocabularies that are significantly different than those represented by either large corpus or pedagogical word lists.

Introduction

Over the last decade there has been active interest in two types of word lists, large corpus word lists and pedagogical word lists. The first is based on a rank ordering of the words in a sample while the second is a limited pedagogical list constructed to represent significant portions of textbook vocabulary. The large corpus word lists are of interest to researchers attempting to understand more about our language. Teachers teach the words on pedagogical lists because they feel they are important for their students' success in reading. One of the most influential pedagogical lists is the Dolch (1936).

Lee Gunderson (PhD, University of California, Berkeley) is an Assistant Professor in the Department of Language Education at the University of British Columbia. His interests are in the area of first and second language reading acquisition.

Dolch produced "A Basic Sight Vocabulary of 220 Words, Comprising All Words, Except Nouns, Common To The Word List of the International Kindergarten Union, The Gates List, And The Wheeler-Howell List." Dolch added some words that were not common to all three lists because:

Many of these words obviously belong with others on the basic list. "Which" belongs with "who" and "that," "done," and "goes" belong with "did" and "go," "start" belongs with "stop," and "write" with "read." (Dolch, 1936, pp. 457, 459)

In 1960 Dolch took several 1000-word samples from reading, math, geography and history texts from grades one to six and found that the words on his list represented a great deal of the vocabulary encountered; for instance, 70 percent of the vocabulary in grade one reading texts to 52 percent of the vocabulary in grade six history texts.

Dolch suggested his list be taught in order to help students develop a sight vocabulary useful in reading all subject areas. He stated, "It is not claimed that the basic list of 220 words includes all the words that the elementary school pupil should know by sight, the claim is only that he should at least know these" (Dolch, 1936, p. 459).

As Hillerich (1974) pointed out, "most educators accept the need for a basic list" (p. 354). For years teachers have used the Dolch as a basic list of sight words to be taught. Occasionally primary teachers teach the Dolch words to their students *before* they involve them in formal reading programs. Some use it as a diagnostic and/or criterion measure since they consider the ability to recognize the 220 words by sight to be a significant indicator of reading achievement. Yet there are critics who maintain that the Dolch is outdated.

Johnson (1971a) stated, "The Dolch List, as a corpus, has outlived its usefulness" (p. 449). Johnson compared the Dolch list with the Kucera and Francis (1967) list which was developed from 500 samples of 15 categories of adult literature. He found that 37 percent (82 words) of the Dolch list were not in the top 220 words on the Kucera-Francis corpus. He concluded that the Dolch list should be replaced by the highest frequency Kucera-Francis words. He responded to the criticism that the Kucera-Francis is based on an adult corpus rather than a children's corpus by stating, "*is it not more important for primary reading materials to reflect what exists in present-day American English, than for sight words lists to reflect what occurs in beginning reading material?*" (p. 457).

Johnson (1971a) constructed his own word list on the basis that items would be: 1) in the top 500 on the Kucera-Francis list and 2) used at least 50 times by kindergarten and first grade students in a study of spontaneous oral language (Murphy, 1957). His list contained 306 words. Hillerich (1974) believed this list "tended to overkill with adult words" (p. 359). Hillerich's own list (1974) consisted of 240 starter words that represented a basic language arts vocabulary. He found that 62 of the Dolch words were not found on his list. He noted that the Dolch may be inappropriate rather than outdated.

Johns (1976a), in a series of studies, developed the Revised Dolch List (RDL). First, he eliminated the 31 words not common to at least three of four lists: 1) the Carroll, Davies, and Richman (1971), (a study of 5,088,721 words occurring in grades three to nine textbooks), 2) the Durr (1973) (the 188 most frequent words in primary library books), 3) the 500 most frequent Kucera and Francis (1967), and 4) the 727 words occurring at a minimum frequency of 50 in the Murphy (1957) study (Johns, 1974). Johns (1976b) added words common to at least three of the

lists mentioned above and produced the RDL of 226 words. Johns (1976a) reported that his list represented significant amounts of the vocabulary of basic readers in four reading series. He also noted that his list represented significantly more of the vocabulary in text than did the Dolch.

The Dolch has its supporters, however. Lowe and Follman (1974) compared the first 150 Dolch words with four other lists and found a high degree of similarity. They suggested the first 150 words can be used "without reservation or limitation" (p. 44). Mangieri and Kahn (1977) analyzed the vocabulary in four basal reading series by taking 3,000 word samples and comparing them to the Dolch. They found the Dolch accounted for 58 to 76 percent of the vocabulary in initial reading texts. Seventy percent of the Dolch words occurred at least once in the basals they studied. Mangieri (1978) showed that significant portions of the vocabulary in four basal reading series were accounted for by the Dolch list (66 to 72 percent).

Thomas (1972) produced a list based on the writing of students in grades one to six in Alberta. He randomly selected about two hundred written products from each grade level resulting in a total of 1,287 compositions and a cumulative 117,878 running words. Thomas' rationale for developing a Canadian list was that the needs of educators, textbook authors, and students would be better served by such a list than by lists developed by American researchers using American subjects. He states, "textbook writers have had to rely heavily upon American vocabulary lists which may not necessarily be meeting the particular regional needs in Canada" (Thomas, 1972, p. 243). He infers that the vocabulary of students in Canada is different than the vocabulary of American students. He concluded, however, that a small number of high frequency words constitutes the majority of the vocabulary in students' writing; differences in vocabulary were apparent in relatively low frequency items.

The purpose of this study was to develop a word list from a small sample of words from a unique population of subjects. It was hypothesized that such a list would be similar to other published lists. If the hypothesis was substantiated, the generated list would account for as much of the vocabulary in initial reading texts as other pedagogical lists. Further, findings would show the similarities between the vocabulary of the subjects and the vocabulary of initial reading texts. If confirmed, the hypothesis would suggest that regional vocabulary needs are not met by simply producing regional vocabulary studies in the traditional manner. Such a finding would suggest that alternate methods of studying vocabulary will have to be developed in order to meet the regional needs in Canada.

Method

Subjects

One objective of the study was to collect vocabulary samples from a population that was obviously unique. With this goal in mind, subjects from a large metropolitan school district in the Western United States were selected randomly from three high schools, four middle schools, and ten elementary schools at grades two, five, eight, and eleven. Each of the participating schools had mean scores in total reading on the Comprehensive Test of Basic Skills (C.T.B.S.) below the 50th percentile. A total of 311 eleventh grade subjects, 321 eighth grade subjects, 309 fifth grade subjects and 268 second grade subjects wrote one or more of the three tasks. A random sample of 20 subjects at each grade level was selected from the 43 percent who had written all three products. The sample consisted of approximately 17 percent

Spanish surname, 24 percent Black, 12 percent Filipino, 20 percent Asian, 18 percent White and 9 percent other (e.g., East Indian). Approximately 22 percent were English as a Second Language (E.S.L.) students. These percentages are close to the population percentages of the schools in the sample and the district at large.

Procedure

Each subject wrote three different essays: a description, a narration, and an argument (writing prompts are shown in the appendix). The assignment was printed on the first page of the test booklet and the teacher read it aloud. Subjects were encouraged to try to spell words correctly, but were given no help. A total of forty minutes was given for each writing task.

Each of the 240 compositions was recorded in permanent computer files. Proper nouns and abbreviations were included as written. Illegible items were ignored. Spelling errors were listed separately, but not included in the recordings. Two error types were excluded; 1) obvious errors (i.e., wuz) and 2) context errors (i.e., "I no you were here."). A program entitled, "A Word Counting and Frequency Analysis Program" (Miller, 1975) was used to count words. In this case a word was defined as any string of characters bounded by blank spaces.

Results

Analysis of the second grade data showed there were 2,011 tokens (total words) and 361 types (different words) for a type/token (T/T) ratio of .18. The fifth grade sample contained 3,178 tokens and 582 types (T/T=.18). The eighth and eleventh grade subjects produced 4,145 and 726 (T/T=.18) and 6,951 and 1,161 (T/T=.17) respectively. When combined, the corpus produced 16,285 tokens and 1,675 types (T/T=.10).

The top 250 words were selected and a list compiled. Since the absolute frequency of words 251 to 253 was the same as the frequency of 250, they were included as well. Inspection of the 253 words showed there were some specifically topic prompted items, i.e., the word *special* in the description task. Eleven words were eliminated as being specifically topic-prompted. The 242 words can be seen in Table 1.

The list contains many inflected forms, i.e., *walk*, *walks*, *walked*, *walking*. Unlike the Dolch list, it also contains nouns such as *mother*, *father*, *birthday*, *teacher*, etc.

Pedagogical List Results

The Inner-City list was compared with four pedagogical lists: Dolch (1936), Johns RDL (1976), Hillerich (1974), and the Durr (1973). The hypothesis that the I.C. list would be similar to other pedagogical lists was confirmed. The I.C. and Dolch lists share 144 words (60 percent of the I.C.), Johns RDL 155 (64 percent), Hillerich 176 (72 percent) and the Durr 146 (60 percent). The I.C. list accounts for 65 percent of the Dolch, 69 percent of the Johns, 73 percent of the Hillerich and 78 percent of the Durr.

The 100 most frequent words on the Thomas (1974) Canadian word list representing 58 percent of his total word count and the I.C. list shared 94 words. The six Thomas words not found on the I.C. list were *dog*, *people*, *two*, *time*, *dad*, and *once*.

TABLE 1
RANK ORDER LIST OF INNER-CITY WORDS

the*	then*	other	talk	say*	again*
to*	get*	has	their*	wanted	an*
and*	what*	into	around*	way	call*
you*	there*	nice	ever	back	end
he*	about*	too*	everything	can't	happy
is*	all*	play*	inside	doesn't	helps
I*	good*	hope*	looked	family	knew
a*	this*	much*	most	keep*	little*
me*	see*	some*	over*	kind*	next
it*	bad*	take*	since	lot	pick*
in*	out*	father	try*	name	problem
my*	will*	something	walking	only*	same
she*	boy	us*	come*	them	seen
was*	just*	got*	didn't	together	still
for*	saw*	could*	found*	did*	sure
of*	up*	how*	house	doing	walk*
that*	they*	even	well*	eat*	won't
your*	or*	help*	work*	goes*	year
his*	had*	no*	after*	last	big*
so*	want*	now*	each	need	class
with*	go*	life	give*	our*	close
have*	please*	might	problems	she's	coming
when*	think*	care	said*	should	door
very*	always*	he's	told	two	down*
but*	school	it's	also	use*	funny*
can*	who*	more	birthday	walked	gets
her*	from*	by*	first*	which*	gives
we*	love	fun	look*	another	left
like*	fine	home	sick	before*	let*
on*	tell*	make*	teacher	brother	living
because*	would*	sometimes	that's	food	made*
not*	day*	came*	years	its*	makes
do*	at*	does*	am*	looks	night
if*	why*	were*	any*	lots	somebody
don't*	man	I'm*	anything	many*	wish
him*	as*	been*	away*	may*	write
are*	things	money	long*	off*	wrong
know*	went*	open	looking	sister	
be*	took	put*	mom	than	
mother	going*	really	never*	thing	
one*	opened	right*	old*	yourself	

* Included on Dolch list

Large Corpus Results

The hypothesis that the I.C. list would be similar to other published lists was substantiated. Correlation coefficients were calculated for the association between I.C. frequencies and frequencies on the Durr (1973), the Carroll, Davies, and

Richman (1971) and the Rinsland (1947), a study of the 6,012,359 words written by students in grades one to eight. The correlations are Durr .83, Rinsland .88, and Carroll, Davies, and Richman .73.

Initial Reading Text Results

The hypothesis that the I.C. list would be as successful as the others in accounting for the vocabulary in initial reading texts was confirmed. The Macmillan (1980) "series r" preprimers and primers (levels 4, 5, 6 and 7) were analyzed. It was found they contained 4,620 words in connected discourse with 323 types for a T/T of .07. Considerable portions of the vocabulary were accounted for by the different lists; 129 I.C. words accounted for 69 percent, 115 Dolch words accounted for 69 percent, 115 Durr words accounted for 68 percent, 117 Johns words accounted for 67 percent, and 120 Hillerich words accounted for 67 percent.

Levels 2, 3, 4 and 5 are the preprimer, primer, and first grade readers in the Ginn (1979) 720 reading series ("Canadianized" version). They contained 11,148 tokens and 549 types (T/T=.05). Sixty-four percent of the vocabulary in the four levels was represented by 139 I.C. words. This compares with the other lists: 1) 139 Dolch words represented 64 percent of the vocabulary, 2) 138 Johns words represented 63 percent of the vocabulary, 3) 133 Durr words represented 65 percent of the vocabulary and 142 Hillerich words represented 60 percent of the vocabulary.

Allyn and Bacon (1977) Pathfinder Series levels 6, 7, 8 and 9 are the first grade readers and contain 11,531 words in connected discourse (745 types, T/T=.06). The I.C. list shares 171 words with the Allyn and Bacon which is 66 percent of the total vocabulary. The Dolch list accounted for 63 percent of the vocabulary in 163 words while the Johns list accounted for 61 percent in 153 words. One hundred fifty-six of the Durr words accounted for 64 percent of the Allyn and Bacon first grade vocabulary. Finally, 163 Hillerich words accounted for 59 percent of the Allyn and Bacon first grade vocabulary. Correlation coefficients were also calculated for the association between I.C. frequencies and those generated by the three reading series. A correlation coefficient of .84 was computed for Macmillan, .81 for Ginn, and .80 for Allyn and Bacon. There is a high degree of association between the frequencies of the words on the I.C. list and in the initial reading texts.

The I.C. list came from a small corpus of words, yet it was as representative of the cumulative vocabulary of the initial reading texts sampled as the other basic lists. The similarities were compelling. An intriguing question was prompted by these findings; would an even smaller sample produce a representative list?

The Second Grade List

The second grade (S.G.) corpus included a total of 2,011 correctly spelled words with 361 different units, representing a mean of approximately 34 (33.50) words per essay. A rank order list of words representing the same proportion of the vocabulary as the I.C. list was compiled. Specifically prompted words were deleted and tied rankings were included. The S.G. list contains 105 words and is shown in Table 2.

Most of the S.G. words are on the lists discussed previously. Eighty-one of them are on the Dolch, 84 on the Johns, 82 on the Durr, 88 on the Hillerich, and 64 on the Thomas 100 most frequent list. An indication of the degree to which the second grade students' vocabulary is different from the older students in the study can be seen by the words on the S.G. list not on the I.C. list. They include such items as

God, time, dad, and better. Other items are clearly artifacts of the assignments, i.e., *box, room,* etc. Seventy-nine of the S.G. words represent 57 percent of the vocabulary in the Macmillan texts, 86 represent 54 percent of the Ginn vocabulary, and 95 represent 53 percent of the Allyn and Bacon vocabulary. The S.G. list accounts for over half of the vocabulary in the initial reading texts.

TABLE 2

RANK ORDER LIST OF SECOND GRADE WORDS

the	that	why
is	very	am
he	want	at
and	but	lots
you	do	no
to	have	now
I	of	said
me	took	time
box	God	when
my	all	again
it	from	anything
a	if	be
in	on	better
she	please	birthday
for	we	dad
bad	came	eat
man	give	end
your	good	father
so	know	going
him	one	had
not	open	how
then	tell	into
don't	there	just
because	this	make
can	too	mom
his	went	nice
like	food	other
saw	gave	put
are	her	say
with	mother	should
room	play	something
will	see	they
boy	tall	think
us	things	what
was	walking	would

The correlations between the frequencies generated by the S.G. words and the large corpus lists are Durr .74, Rinsland .77, and Carroll, Davies, and Richman .64. The ordering of language appears to be a robust characteristic varying little from sample to sample. The unique items of the vocabulary of a particular population

are evident in relatively low frequency items. One more question was explored: does pooling the vocabulary of textbooks mask the uniqueness of individual texts?

The lowest level textbook from each series was analyzed. The Allyn and Bacon level six contained 1,836 words, the top ten representing 36 percent of the volume's vocabulary—*the, is, big, on, a, I, am, Ann, Nat, and Nate*. Forty-five percent of the series "r" level four vocabulary was represented by the top ten words—*I, the, and, why, to, jump, he, like, does, and says*. In descending order the ten words *Bill, and, Jill, Lad, is, can, here, this, Nan, and Ted* represented 55 percent of Ginn 720 level two. The T/T ratios for the three texts were, respectively, .07, .13, and .02. These initial texts varied in the rate of repetition of vocabulary (a low T/T indicates high repetition). The I.C. list accounted for 48 percent of the Ginn, 64 percent of the Macmillan, and 57 percent of the Allyn and Bacon vocabulary. Portions of the vocabulary accounted for by the other lists are: 1) Ginn—Dolch 44 percent, Durr 54 percent, Johns 54 percent, Hillerich 46 percent; 2) Macmillan—Dolch 71 percent, Durr 65 percent, Johns 64 percent, Hillerich 59 percent; 3) Allyn and Bacon—Dolch 50 percent, Durr 53 percent, Johns 54 percent, Hillerich 55 percent. Seventeen percent of the Macmillan vocabulary was not included on any of the lists (two percent were proper nouns). Thirty-five percent of the Ginn vocabulary did not occur on any of the lists (30 percent were proper nouns). Thirty percent of the Allyn and Bacon vocabulary was not found on any of the lists (17 percent were proper nouns). The S.G. list, with considerably fewer words accounted for 45 percent of the Macmillan, 43 percent of the Ginn, and 44 percent of the Allyn and Bacon vocabularies.

Discussion

The I.C. word list accounts for a substantial portion of the vocabulary in the initial reading texts analyzed. The percentage of vocabulary the I.C. list represents is as high or higher in all cases as the percentage represented by the Dolch list, the Durr list, the Hillerich list, and the Johns RDL.

The similarity of the ordering of words between the I.C. and other vocabulary lists is compelling. The I.C. ordering is most similar to the Rinsland list which is also based on written material. It is also similar to the vocabulary in children's trade books, although to a lesser degree. It is least similar to the vocabulary of textbooks used in grades three to nine (Carroll, Davies, & Richman).

The similarity between the I.C. list and Thomas' Canadian list is compelling. Fifty-eight percent of all the words in the Thomas list were accounted for by one hundred items, the I.C. list containing 94 of them. The similarities between the vocabularies of Canadian students and inner-city American students are striking.

The I.C. list was developed from a relatively small corpus of words, yet it appears to be representative when compared with other lists and with the vocabulary of initial reading texts. A list was developed from an even smaller corpus of words, the second grade essays. The S.G. list of 105 words accounts for 53 to 57 percent of the vocabulary in the initial reading texts, remarkable considering the average size of the essays was approximately 34 words. It appears the relative frequencies of words (and particularly function words) remains fairly constant from sample to sample, supporting Zipf's contention that the distribution of word frequencies occurring in any *ad libitum* sample of language always has the same distribution (Zipf, 1935). He notes,

Perhaps the most interesting feature of this high degree of orderliness in the distribution of words in the stream of speech is this: we select and arrange our words according to their meanings with little or no conscious reference to the relative frequency of occurrence of those words in the stream of speech, yet we find that words thus selected and arranged have a frequency distribution of great orderliness which for a large portion of the curve seems to be constant for language in general. (p. 48)

The writing samples were typical classroom assignments and so were not *ad lib*, but when text-prompted items were removed the list became relatively representative. The vocabulary of individual initial reading texts is not *ad libitum*. Two types of texts appeared. The first was characterized by the use of names and repetition while the second was characterized by the use of pronouns and verbs. The word lists accounted for considerably more of the vocabulary in the second type of text than in the first. The use of names and repetition was a feature of the initial reading texts of the forties and fifties. Have we given Dick and Jane new names?

Word lists are successful in representing substantial portions of the vocabulary in reading texts when the vocabulary of the reading texts is analyzed as a combined sample. To ascertain the vocabulary load for a particular text, one must analyze that text. It appears that pedagogical lists are more accurate in preparing students for cumulative vocabulary than they are for the vocabulary of individual texts.

Conclusion

Teachers use word lists as sources of words to be taught to students as sight words. Authors attempt to compile lists that represent large portions of the vocabulary encountered in school texts. The general belief is that students who are able to recognize important sight words will have command of the majority of words they encounter in text. While this does not guarantee they will comprehend the text, many teachers believe it will aid comprehension since less time and effort need to be directed to recognition activities.

Findings of the present study show that a word list can be developed from a small language sample of a small number of subjects. The I.C. word list is as successful in accounting for the cumulative vocabulary in initial reading texts as are other lists.

It was hypothesized that the I.C. list would account for much of the vocabulary of initial reading texts. This hypothesis was confirmed. There are many published word lists and, as shown in this paper, they represent considerable portions of the printed text students will cumulatively encounter. The I.C. list is as successful as several other well-known lists in representing the vocabulary of initial reading texts. Hillerich (1974) notes, "wars may be *escalated*, and people *fragged* or *televised* while few sit in a *parlor* anymore, but we have yet to replace or supplement *the*, *and*, or *to*." This would seem true for the written language of students in the inner-city where "standard English" is not the mode. Indeed, it would seem to be true for students in different regions across Canada and the United States.

Pedagogical lists are most often compiled by consulting other lists that are derived from large samples of words. As seen here, this is not necessary. Indeed, the unique character of the vocabulary of subjects and individual textbooks is lost when the sample size is large. This phenomenon is particularly compelling in the case of initial reading texts. As far as pedagogical lists are concerned, they all represent significant portions of the cumulative vocabulary in initial reading texts, somewhere between 60 and 70 percent. So any major pedagogical list will serve us equally well.

We should begin, however, to look at the vocabulary not represented by the major lists, the 30 to 40 percent of the vocabulary that is not included. To meet the regional needs of a particular area we must analyze the relatively low frequency items of vocabulary that represent its unique character. In this manner the distinctively Canadian vocabulary can be identified in various geographic areas. The availability of personal computers to the classroom makes such a task possible for teachers. As far as pedagogical lists are concerned, we have done enough.

References

- Allyn & Bacon. (1977). *Pathfinder. Allyn and Bacon Reading Program*. Boston: Author.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book*. Boston: Houghton Mifflin; New York: American Heritage.
- Dolch, E. W. (1936, February). A basic sight vocabulary. *Elementary School Journal*, 36, 456-460.
- Dolch, E. W. (1960). *Teaching primary reading*. Champaign, IL: Garrard.
- Durr, W. K. (1973). A computer study of high-frequency words in popular trade juveniles. *The Reading Teacher*, 27(14), 37-42.
- Ginn and Company. (1979). *Reading 720*. Scarborough, ON.
- Hillerich, R. L. (1974, January). Word lists—getting it all together. *The Reading Teacher*, 27(1), 353-360.
- Johns, J. L. (1974). Should the Dolch list be retired, replaced, or revised? *Elementary School Journal*, 74, 375-380.
- Johns, J. L. (1976a). Updating the Dolch basic sight vocabulary. *Reading Horizons*, 16, 104-111.
- Johns, J. L. (1976b). Some comparisons between the Dolch basic sight vocabulary and the word list for the 1970s. *Reading World*, 15(3), 144-150.
- Johnson, D. (1971a, February). The Dolch list reexamined. *The Reading Teacher*, 24(5), 449-457.
- Johnson, D. (1971b, October). A basic vocabulary for beginning reading. *Elementary School Journal*, 72, 29-34.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Lowe, A. J., & Follman, J. (1974, October). Comparisons of the Dolch list with other word lists. *The Reading Teacher*, 28(1), 40-44.
- Macmillan Publishing Company. (1980). *Series r, Macmillan reading*. New York: Author.
- Mangieri, J. N., & Kahn, M. S. (1977, March). Is the Dolch list of 220 basic sight words irrelevant? *The Reading Teacher*, 30(6), 649-651.
- Mangieri, J. N. (1978, October). Dolch list revisited. *Reading World*, 18(1), 91-95.
- Miller, A. (1975). *A word counting and frequency analysis program*. Vancouver, BC: University of British Columbia Computing Centre.
- Murphy, H. A. (1957, December). The spontaneous speaking vocabulary of children in the primary grades. *Journal of Education*, 140(2), 2-105.
- Rinsland, H. D. (1947). *A basic vocabulary of elementary school children*. New York: The Macmillan Co.
- Thomas, V. (1972, December). The basic writing vocabulary of elementary school children. *The Alberta Journal of Educational Research*, 18(4), 243-248.
- Thomas, V. (1974). *Teaching spelling Canadian word lists and instructional techniques*. Toronto: Gage.

Zipf, G. K. (1965). *The psycho-biology of language, an introduction to dynamic philology*. Cambridge, MA: M.I.T. Press (original work published in 1935).

APPENDIX

Description Prompt:

We all know someone who is special. People can be special in many different ways. Describe a person who is special to you. Tell why you think he or she is special. If you do not know how to spell a word, just spell it the way you think it sounds. Since we do not care about misspelled words, please do not ask for help in spelling.

Argument Prompt:

Pretend your best friend has just started smoking for the first time and you have decided to write him a letter. Since you know smoking is not good for your friend, you must try to get him to stop. You are going to do this by writing him a letter. If you do not know how to spell a word, just spell it the way it sounds. Since we do not care about misspelled words, please do not ask for help in spelling.

Narration Prompt:

Look carefully at the three pictures. Please write a story about these three pictures. Tell what is going on in each picture. What do you think is happening? Remember, the three pictures go together so make one story. If you do not know how to spell a word, just spell it the way you think it sounds. Since we do not care about misspelled words please do not ask for help in spelling.



STEPHEN T. CAREY

University of Alberta

and

JAMES CUMMINS

Ontario Institute for Studies in Education

Communication Skills in Immersion Programs

A study of cognitive and behavioral variables, academic achievement, and English and French speaking skills was undertaken to find the correlates of high French oral communication skills of grade three francophone and anglophone immersion students as rated by teachers and independent francophone raters. Teachers' ratings correlated with academic achievement and cooperativeness variables whereas those of independent francophone raters did not. The need for further research and the development of objective procedures for the evaluation of oral communications skills is discussed in terms of the influence that such evaluations might have on students' attitudes and hence the success of immersion programs.

One of the concerns related to the evaluation of immersion programs is the lack of adequate assessment of oral production and aural comprehension skills that have been performed for these programs. This deficiency exists in part due to the time constraints and monetary restrictions that exist for most evaluation programs, as well as the difficulty in finding qualified oral examiners.

Nevertheless, since one of the prime goals of immersion programs is to develop functional oral communication skills, it is important that evaluation techniques be developed, refined and understood in terms of their ecological validity. The present study sought to investigate what variables correlated with teachers' ratings of their students' oral French skills and how these correlations between variables and ratings compared with similar ratings by independent francophone teachers. A com-

Stephen T. Carey obtained his doctorate from the University of Toronto and teaches at the Faculté Saint-Jean of the University of Alberta. His research interests include memory and bilingualism.

James Cummins obtained his doctorate from the University of Alberta and teaches at the Modern Language Centre of the Ontario Institute for Studies in Education. His research interests center on cognitive aspects of bilingualism and language acquisition.

parison of these sets of ratings, it was believed, might shed light on the process of evaluation of students' oral skills by their teachers.

An additional concern of researchers, teachers, and parents is to identify the cognitive, behavioral, and background variables that are associated with students who do well in immersion programs, in order that some objective basis might be established for predicting if a given student will or will not do well in an immersion program. Such information would be useful in counseling parents with regard to placing or withdrawing their children from an immersion program. The present study, therefore, sought to determine what cognitive, behavioral, and background variables are associated with students who do well in immersion programs to gain further insight into the dynamics of these programs.

Students in grade three were examined in this study rather than children in the later grades for several reasons. First, the authors believe that students' initial progress in acquiring French and English communication skills is instrumental in determining their attitude towards the immersion program. Second, the authors believe that the early establishment of a positive attitude towards the French immersion program is instrumental to later success in the program. Thus, it is suggested that the students' perception of their initial success may be critical in determining their subsequent achievement in this domain, and inasmuch as the child's perceptions of his success is influenced by the teacher's perception of his competence, a further understanding of the basis of teacher's evaluations is critical.

Language achievement testing at this level is problematic and the authors chose to rely on francophone teachers and independent francophone testers for providing criterion measures of students' oral French achievement. In addition, a study of the relationship between teachers' ratings of francophone and anglophone students' oral French achievement and the ratings of independent francophones of the same students, it was believed, might provide insights into how teachers perceived each of these student groups. Finally, the authors believe that a study of the relationship between cognitive, behavioral, and background variables of the students and the evaluation of their French achievement by the two rating groups might reveal differences in teachers' perception of the students and also lead to a description of students who are likely to achieve well in immersion programs.

Subjects

The subjects were 43 children drawn from all of the grade three classes of three Edmonton Catholic Separate School (ECSS) bilingual schools. Teachers were asked to specify the linguistic background of their pupils as either anglophone or francophone. Each child was administered the Raven's Progressive Matrices (a nonverbal intelligence test). The 43 children in the study were selected on the basis of having anglophone and francophone groups who were equivalent in terms of Raven's scores. However, when the children were classified according to their responses on the linguistic background questionnaire (Appendix C), it was found that many of those students classified by teachers as coming from English backgrounds, actually came from homes where a considerable amount of French was spoken (at least by parents to children). Consequently, the groups were classified into English, French, and mixed French-English home/backgrounds. According to this classification, 24 children came from English homes, 10 from predominantly French and 9 from mixed French-English homes.

Instruments

The Teacher Rating and students' questionnaires were the same as those used in the Carey and Cummins (1983) grade five study and are presented in Appendices B and C. Based on children's responses to items 1-4 on the student questionnaire, an index of overall home French usage was constructed where 1 indicated English only spoken, 2 indicated some French spoken, 3, a considerable amount of French spoken, and 4, French only spoken. Groups three and four formed the francophone group and group two the mixed French-English group. A rating of two signified that at least one parent spoke French most often to the child while a rating of three signified that either both parents spoke French most often to the child or that the child interacted most often in French with at least one parent (i.e., both parent and child speak mostly French to one another). This variable is number 15 in Table 1.

English achievement was assessed by means of the Gates McGinitie Reading Test and Vocabulary and Comprehension Scales, administered by the ECSS at the grade two level. A measure of *English* syntax comprehension developed by Tremaine (1975) was also administered. The Primary Mental Abilities (PMA) IQ test, administered by the ECSS at the grade one level, was also included in the analysis.

French skills were assessed in three ways: first, by means of a *French* syntax comprehension test developed by Tremaine (1975). Five syntactical constructions were assessed for both French and English syntax comprehension. English and French versions were translations of one another, and the French version was administered first to all subjects (see Appendix A for examples of test items). Approximately two weeks separated administration of English and French versions. Each of the sentences was read to the subjects who were required to pick one of three pictures which illustrated the correct meaning of the sentence. Four sentence exemplars were read for each of the five syntax construction types. Order of sentence presentation was random. The five constructions were as follows (from easiest to most difficult) according to the norms supplied by Tremaine (1975).

1. *Relative Clause*
The baby eats the cake which is on the table.
The baby on the table eats the cake.
2. *Passive*
The baby is seen by the girl.
The girl is seen by the baby.
3. *Reflexive*
The boys see themselves.
The boys see each other.
4. *Direct Object*
The girl shows the cow to the dog.
The girl shows the dog to the cow.
5. *For-to*
The baby gives the ball for the dog to the cat.
The baby gives the ball for the cat to the dog.

Teachers were also asked to rate children's proficiency in French on four five-point scales: French reading vocabulary, French reading comprehension, French oral expression, and French aural comprehension. These four measures correlated very highly (range .65-.94) with each other indicating that teachers were possibly in-

advertently giving a global rating of French proficiency rather than discriminating between different subskills. Consequently, a combined score from the four scales was used in the analysis.

TABLE 1

Variable		Teacher French Rating	French Speaking Rating (Testers)
1.	Teacher French Rating	1.00	.23
2.	French Speaking Rating	.23	1.23
3.	Gates Vocabulary	.44*	-.41*
4.	Gates Comprehension	.51*	-.23
5.	Raven's	.41*	.16
6.	PMA	.39*	-.29
7.	Language Spoken by Mother (SQ1)	-.08	.47*
8.	Language Spoken by Father (SQ2)	.22	.22
9.	Language Spoken to Mother (SQ3)	.23	.27
10.	Language Spoken to Father (SQ4)	.22	.07
11.	Language Spoken to Siblings (SQ5)	.07	.05
12.	Happy to Speak French Elsewhere (SQ6)	.51*	.09
13.	Language Preference to Teacher (SQ7)	.51*	.07
14.	Language Preference to Friends (SQ8)	.06	-.10
15.	Overall Home French Usage	.19	.37*
16.	English Syntax (Es) Total	.38*	.20
17.	Es Relative	.33*	.08
18.	Es Passive	.19	-.32*
19.	Es Referential	.29	-.02
20.	Es Direct Object	.28	.36*
21.	Es For-To	.15	.04
22.	French Syntax (Fs) Total	.54*	.17
23.	Fs Relative	.30*	.06
24.	Fs Passive	.40*	-.04
25.	Fs Referential	.24	-.19
26.	Fs Direct Object	.09	-.04
27.	Fs For-To	.45*	.35*
28.	Asks Questions (TR1)	.41*	.03
29.	Responds to Questions (TR2)	.61*	.14
30.	Group Participation (TR3)	.58*	.10
31.	Friends (TR4)	.59*	-.14
32.	Lively (TR5)	.52*	-.24
33.	French at Recess (TR6)	.61*	.20
34.	Asks Teacher Help (TR7)	.26	-.22
35.	French in Corridors (TR8)	.56*	.19
36.	Extra French at Home (TR9)	.46*	.24
37.	Sex (2 = male, 1 = female)	.35*	.23

* = $p \leq .05$

** = $p \leq .01$

Variable		Gates Vocabulary	Gates Comprehension
1.	Teacher French Rating	.44*	.51*
2.	French Speaking Rating	-.41*	-.23
3.	Gates Vocabulary	1.00	.81
4.	Gates Comprehension	.81*	1.00
5.	Raven's	.21	.40*
6.	PMA	.71*	.69*
7.	Language Spoken by Mother (SQ1)	-.47*	-.53*
8.	Language Spoken by Father (SQ2)	-.19	-.22
9.	Language Spoken to Mother (SQ3)	-.16	-.24
10.	Language Spoken to Father (SQ4)	-.01	-.06
11.	Language Spoken to Siblings (SQ5)	-.22	-.19
12.	Happy to Speak French Elsewhere (SQ6)	-.16	-.34*
13.	Language Preference to Teacher (SQ7)	.38*	.41*
14.	Language Preference to Friends (SQ8)	.17	.03
15.	Overall Home French Usage	-.23	-.30*
16.	English Syntax (Es) Total	.35*	.51*
17.	Es Relative	.44*	.50*
18.	Es Passive	.38*	.43*
19.	Es Referential	.13	.17
20.	Es Direct Object	.13	.34*
21.	Es For-To	.14	.11
22.	French Syntax (Fs) Total	.37*	.48*
23.	Fs Relative	.27	.12
24.	Fs Passive	.26	.45*
25.	Fs Referential	.16	.12
26.	Fs Direct Object	.24	.42*
27.	Fs For-To	.29	.27
28.	Asks Questions (TR1)	.31*	.39*
29.	Responds to Questions (TR2)	.30*	.44*
30.	Group Participation (TR3)	.27	.46*
31.	Friends (TR4)	.38*	.54*
32.	Lively (TR5)	.39*	.48*
33.	French at Recess (TR6)	.32	.41*
34.	Asks Teacher Help (TR7)	.10	.26
35.	French in Corridors (TR8)	.10	.27
36.	Extra French at Home (TR9)	.11	.17
37.	Sex (2= male, 1=female)	.19	.10

* = $p \leq .05$ ** = $p \leq .01$

The third measure of French proficiency was a rating of French speech production made by a native francophone teacher on the basis of a 10-15 minute interview with each child. The interview consisted of casual conversation related to the child's interests, as well as activities such as describing a set of cartoon pictures and listening to and retelling a French story. The rating was made on a 7-point scale, ranging from "speaks French like a native speaker of his/her age" to "has almost no command of spoken French."

Results

The correlations of the variables with French and English language skills are presented in Tables 1 and 2 and differences between the three groups of students are presented in Table 3.

TABLE 2

Variable		French Syntax Total	English Syntax Total
1.	Teacher French Rating	.54*	.38*
2.	French Speaking Rating	.17	.20
3.	Gates Vocabulary	.37*	.35*
4.	Gates Comprehension	.48*	.51*
5.	Raven's	.37*	.34*
6.	PMA	.51*	.48*
7.	Language Spoken by Mother (SQ1)	.07	-.13
8.	Language Spoken by Father (SQ2)	-.08	-.06
9.	Language Spoken to Mother (SQ3)	.25	-.05
10.	Language Spoken to Father (SQ4)	.04	-.03
11.	Language Spoken to Siblings (SQ5)	.20	-.01
12.	Happy to Speak French Elsewhere (SQ6)	.27	.36*
13.	Language Preference to Teacher (SQ7)	.57*	.25
14.	Language Preference to Friends (SQ8)	.09	.13
15.	Overall Home French Usage	.12	-.08
16.	English Syntax (Es) Total	.66*	1.00
17.	Es Relative	.37*	.64*
18.	Es Passive	.24	.40*
19.	Es Referential	.33*	.50*
20.	Es Direct Object	.49*	.65*
21.	Es For-To	.33*	.47*
22.	French Syntax (Fs) Total	1.00	.66*
23.	Fs Relative	.43*	.19
24.	Fs Passive	.50*	.35*
25.	Fs Referential	.49*	.16
26.	Fs Direct Object	.47*	.52*
27.	Fs For-To	.70*	.47*
28.	Asks Questions (TR1)	.49*	.28
29.	Responds to Questions (TR2)	.54*	.37*
30.	Group Participation (TR3)	.48*	.33*
31.	Friends (TR4)	.25	.18
32.	Lively (TR5)	.19	.11
33.	French at Recess (TR6)	.46*	.30*
34.	Asks Teacher Help (TR7)	.25	.07
35.	French in Corridors (TR8)	.49*	.30*
36.	Extra French at Home (TR9)	.38*	.29
37.	Sex (2= male, 1= female)	.30*	.07

* = $p \leq .05$

** = $p \leq .01$

TABLE 3
T-TESTS FOR DIFFERENCES BETWEEN GRADE 3 GROUPS

	Anglophone	(N=24)	Means Francophone	(N=10)	Mixed	(N=9)
1.	14.3		16.00		12.8	
2.	41.8	**	51.0	**	36.1	
3.	36.7		29.1		38.3	
4.	39.4	**	28.1		33.5	
5.	28.6		29.6		25.6	
6.	114.4		108.2		112.8	
7.	2.0	**	3.8	**	2.7	
8.	2.0	**	3.6	**	2.3	
9.	2.0	**	3.4	*	2.4	
10.	2.0	**	3.1	*	2.2	
11.	2.0	**	2.7		2.3	
12.	3.0	*	3.6	*	2.6	
13.	2.5		2.6		2.6	
14.	1.4		1.6		1.4	
15.	---		---		---	
16.	13.3		12.5		13.0	
17.	3.3		3.0		3.1	
18.	3.3	**	2.4	*	3.3	
19.	2.5		2.4		2.8	
20.	3.0	*	2.8		2.3	
21.	1.2		1.9		1.4	
22.	13.8		14.0		14.0	
23.	3.2		3.5		3.8	
24.	3.4		3.3		3.3	
25.	2.5		2.0		2.7	
26.	3.1		2.7		3.0	
27.	1.6		2.4		1.2	
28.	4.0		3.8		4.4	
29.	4.7		4.0		4.1	
30.	4.6		4.3		4.4	
31.	4.5		4.0		4.1	
32.	4.5		4.3		4.3	
33.	2.3		3.1		2.0	
34.	3.8		4.0		4.1	
35.	2.1	*	2.9		2.0	
36.	---		---		---	

* = $p \leq .05$

** = $p \leq .01$

Correlates of French Skills

It is clear from the pattern of correlations that Teacher French Rating (TFR) and French Syntax Total (FST) are indicative of very different dimensions of proficiency as compared to French Speaking Rating by an independent rater (FSR). Tables 1 and 2 show that TFR and FST both correlate highly with measures of academic achievement or potential (e.g., Gates Vocabulary and Comprehension,

Raven's, PMA, English Syntax). Thus, TFR and FST together with these English measures appear to be assessing a dimension of academic proficiency which is independent of language of testing, language spoken at home, or French speech rating by an independent francophone rater. Thus, teachers rate students who score high on cooperativeness and academic success and English language skills as also being proficient in French. In contrast, native francophone raters' judgments of the students' oral French proficiency did not show these correlations.

Teacher French ratings were significantly related to most of the indices of teachers' ratings of students' behavior (variables 28-36). These variables appear to be tapping a dimension of academic cooperativeness. Thus, these relationships are consistent with our interpretation of the teachers rating academic and English competence, cooperativeness, and French as a unitary dimension. The strong relationships between TFR and FST and children's language preference to teacher (variable 13) is also consistent with this pattern.

French Speech Ratings (FSR) by independent raters present a different picture. They are positively related to the use of French at home, especially language spoken by the child's mother (variable 7) but tend to be less highly correlated with English academic skills than do teachers' French ratings. This latter finding can be accounted for by the fact that the francophone group tends to do more poorly on measures of English academic and cognitive skills (variable 4) than the anglophone group (see Table 3). The lack of significant relationships between FSR and intelligence and academic French skills is consistent with Genesee's (1976) finding that the development of communication skills in French immersion programs is independent of children's intellectual ability. These results also lend support to the proposed distinction between Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP) (Cummins, 1979; Swain, 1981).

Correlates of English Skills

The patterns of correlations for English academic skills are almost identical to those for French academic skills—compare, for example, the correlations for English and French syntax and for Gates Comprehension and Teacher French Rating. The only group of variables which do not conform to this pattern are the indices of home language use and children's willingness to speak French (variable 12) which are negatively related to Gates English Comprehension probably because of the lower scores of the francophone group on this variable.

Group Differences

Because of the small numbers of students in the francophone and mixed groups, group differences must be interpreted cautiously. However, there are some interesting trends in the data. Since the home language background questionnaire was the basis for group selection, it is not surprising that there are group differences on variables 7-11. The francophone group also tends to be more willing to use French outside of school (variable 12) but there are no differences in French preference to teachers and friends among francophone, anglophone, and mixed groups (variables 13 and 14). The francophone group performed more poorly than the anglophone group on Gates Comprehension, but better than the anglophone and mixed groups on French Speaking Rating. They performed more poorly than both the other groups in their comprehension of passives in the English Syntax Test (variable 18) while the mixed group performed worse than the anglophone group in the under-

standing of direct and indirect objects (variable 20). The only other difference was that the francophone group was rated by teacher to use more French in the corridors (variable 35) as compared to the anglophone group.

Discussion

The fact that the francophone group did not perform as well as the anglophone group on Gates Comprehension suggests that at the time when the test was administered (grade two), the francophone group had not yet caught up with the anglophone group in English academic skills. However, other data suggests that this has been accomplished by the grade five level (Carey & Cummins, 1983).

Two additional aspects of the findings seem worthy of special attention. First is the unitary nature of academic skills regardless of the specific language (English or French) of instruction; this means that students who do well in English programs will also do well in French immersion programs and that students who are not performing well in French immersion programs will not necessarily improve if they are transferred into a regular program. Secondly, the acquisition of French speaking skills in an immersion context as rated by independent native French speaking teachers, appears not to be dependent on general intellectual or academic ability. These findings are consistent with findings of previous research on French immersion programs. Genesee (1976) states that the acquisition of French literacy skills is related to the student's IQ level, but the acquisition of interpersonal communication skills is not. Similarly, Carey and Cummins (1983) reported that written performance on French cloze tests correlated highly with measures of intellectual and English academic skills, but not with oral skills.

The consistency of the results of the present study in conjunction with the pattern observed in previous research, warrants examining potential implications for predicting students' academic success. The findings suggest that the same factors underlie the development of academic skills in a French immersion program as in a regular English program. Thus, children who experience *academic* problems in a French program are likely to experience similar academic problems in an English program, and students who do well in unilingual English academic studies will also tend to do well in studies taught in French. A corollary of this is that academic problems will not necessarily disappear when a child who is experiencing difficulties in a French immersion program is switched to an English program. There is some evidence to support this hypothesis (Bruck, 1978, 1982). However, due to the paucity of research regarding the relationship between learning disabilities and success in immersion programs and the existence of contradictory data (Trites, 1976, 1978), no firm conclusions can be drawn at the present time.

The converse implication is that the development of oral communication skills appears to be independent of general academic and intellectual skills and that many children who experience academic difficulties in a French immersion program will nevertheless develop reasonable fluency in French. Future research should delineate the variables that are not associated with success in written French acquisition but are prerequisites for success in oral communication skills. Extroversion and social confidence would appear to be important assets in such immersion pedagogy since an emphasis is placed on oral production in a group setting. Indeed, there is some evidence (Bruck, 1982) that it is frequently the child who is reticent, shy and introverted that may experience difficulty in early immersion programs due to the requirements of oral production in a social group setting.

Finally, the inconsistencies in evaluating oral French competence between teachers and independent francophone raters is extreme in some cases. For example, the correlation between teachers' ratings of French and variable 3, Gates Vocabulary is +44, yet for independent French raters this correlation is -41. Similar reversals are found between the two rating groups for variable 4, Gates Comprehension, variable 5, Raven's Progressive Matrices and variable 6, Primary Mental Abilities. Clearly teachers' ratings of students' *French* achievement are highly correlated with English competence and general academic ability whereas the independent raters' evaluations show that proficiency in oral French communication is independent of such abilities. There is one major implication of this finding. The results demonstrate the need to include several independent measures of oral communication skills in any assessment research design. The issue of how to interpret the results is less clear cut. It appears that the teachers' assessment of the child's oral fluency is biased by their knowledge of the child's academic proficiency and the basic type of interaction they have with the child (i.e., whether the child is cooperative or non-cooperative), and this knowledge masks the teacher's ability to assess French oral expression as a discrete skill.

As previously mentioned, one of the major goals of immersion programs is to promote fluent oral communication skills. This research signals a caution in the teacher's ability to recognize this discrete skill and suggests the need for the availability of alternative objective methods, instruments and procedures for evaluating oral and written French skills in order to overcome the teacher's tendency to rate students' French competence in terms of their academic ability. This would seem to be particularly important for early grades in the immersion program since the student's self-perception and future attitude toward the immersion program are surely influenced by the teacher's evaluation of the student's French communication skills, as well as the teacher's perceptions of the student's language competence and his/her consequent behavioral correlates.

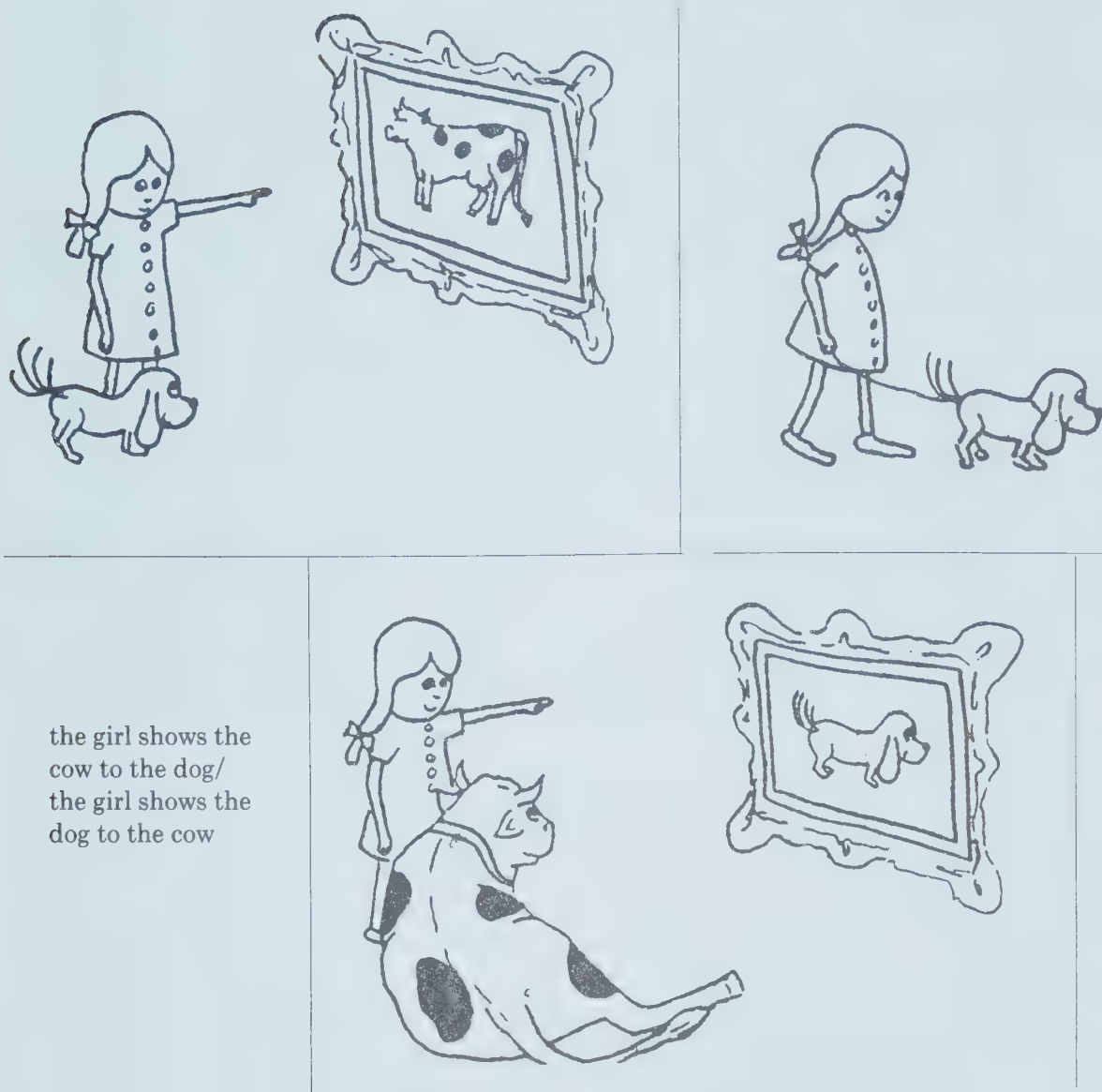
Finally, the implications for teacher preparation programs for immersion teachers would include a consideration of the critical role that students' attitudes play in determining their success in that program and the important role that teachers' evaluations must surely play in influencing the students' attitudes towards the immersion program. Clearly, the assessment of oral communication skills remains highly problematic (Canale, 1982); nevertheless the reliance on the communicative method of instruction in immersion programs requires that teachers have access to methodologies and procedures that can aid them in performing both formative and summative evaluation of oral communication skills. The shift in emphasis from language form to language use that is the hallmark of immersion programs aimed at producing oral communicative performance must be accompanied by an equal shift in the evaluation of communication skills.

Appendix A

English Syntax Test: Passives

	
the boy hits the ball/ the ball hits the boy the boy is hit by the ball/the ball is hit by the boy	

English Syntax Test: Direct/Indirect Object



the girl shows the
cow to the dog/
the girl shows the
dog to the cow

Appendix B

Nom de l'Elève _____

En le comparant avec les autres enfants de votre classe et avec les autres enfants du même niveau que vous avez connus, est-ce que cet enfant:

1.	Pose beaucoup de questions en classe.	_____	Pose rarement de questions en classe.
2.	Répond souvent aux questions posées en classe.	_____	Répond rarement aux questions posées en classe.
3.	Participe volontiers aux activités de groupe.	_____	Ne participe pas aux activités de groupe.
4.	A beaucoup d'amis.	_____	A très peu d'amis.
5.	Est volubile et plein de vie.	_____	Est tranquille, timide et à l'écart.
6.	En récréation s'exprime toujours en français avec ses camarades.	_____	En récréation ne s'exprime jamais en français avec ses camarades.
7.	Recherche l'aide et les conseils du professeur.	_____	Travaille la plupart du temps seul/e.
8.	S'exprime toujours en français dans les couloirs.	_____	Ne s'exprime jamais en français dans les couloirs.
9.	Fait du travail supplémentaire en français à la maison ou à l'école.	_____	Ne fait jamais du travail supplémentaire en français.

Appendix C

Name _____

School _____

- Which language does your *mother* speak at home most of the time?
French _____ English _____ Other _____
French and English about the same _____
- Which language does your *father* speak at home most of the time?
French _____ English _____ Other _____
French and English about the same _____
- Which language do *you* speak with your mother most of the time?
Mostly French _____ Mostly English _____ Other _____
French and English about the same _____
- Which language do *you* speak with your father most of the time?
Mostly French _____ Mostly English _____ Other _____
French and English about the same _____
- Which language do *you* speak with your brothers and sisters most of the time?
Mostly French _____ Mostly English _____ Other _____
French and English about the same _____
- How happy or unhappy are you to meet people outside of school whom you can speak French to?
Very happy _____ Fairly happy _____ Fairly unhappy _____
Very unhappy _____
- In school, do you prefer to speak English or French to your teacher?
English _____ French _____ Both about the same _____
- In school, do you prefer to speak English or French to your friends?
English _____ French _____ Both about the same _____

References

- Bruck, M. (1978). The suitability of early French immersion programs for the language disabled child. *Canadian Journal of Education*, 3, 51-72.
- Bruck, M. (1982). Language impaired children's performance in an additive bilingual education program. *Applied Psycholinguistics*, 3, 45-60.
- Canale, M. (1981). From communicative competence to communicative language pedagogy. In J. C. Richards & R. Schmidt (Eds.), *Language and communication*. London: Longman.
- Carey, S. T., & Cummins, J. (1983). Achievement, behavioral correlates and teachers' perceptions of francophone and anglophone immersion students. *Alberta Journal of Educational Research*, 29, 159-167.
- Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, 19, 197-205.
- Genesee, F. (1976). The role of intelligence in second language learning. *Language Learning*, 26, 267-280.
- Swain, M. (1981). Time and timing in bilingual education. *Language Learning*, 3, 1-16.
- Tremaine, R. V. (1975). *Syntax and Piagetian operational thought*. Washington, DC: Georgetown University Press.
- Trites, R. L. (1976). Children with learning difficulties in primary French immersion. *Canadian Modern Language Review*, 33, 193-216.
- Trites, R. L. (1978). Learning disabilities in immersion. *Canadian Modern Language Review*, 34, 888-889.

CLAIRE F. STAAB

University of British Columbia

Student Teachers' Perceptions of Their Competence Versus the Perceptions of University Supervisors and Sponsor Teachers

The primary purpose of this study was to determine whether student teachers, their university supervisors, and their sponsor teachers judged the student teacher's performance at the same level of competence over four clusters of teaching behaviors (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities).

Sixty-six student teachers, their university supervisors, and their sponsor teachers were asked to respond to identical questionnaires, rating the student teacher's performance on the above behaviors. An item analysis using Lertap was done to determine the Hoyt estimate of reliability and each of the four clusters was subjected to a one way analysis of variance across three groups of subjects. Pearson correlations were also done to determine the degree of similarity of opinion both between and within groups. Results indicated a significant difference in Preparation for Instruction. Sponsor teachers rated student teachers significantly higher than university supervisors.

Many articles over the past fifteen years have dealt with the notion that communication between student teachers, university supervisors and sponsor teachers needs to improve if desired changes are to be brought about in the student teacher's classroom performance (Larson & Ringness, 1965; Campbell & Williamson, 1973; Reynolds, 1977; Seiferth & Solliday, 1980; Queen & Mallen, 1982). However, little work has been done to determine whether the individual student teacher, his/her university supervisor, and his/her sponsor teacher view that student teacher's performance as being at the same level of competence and thus have a basis for beginning a constructive dialogue regarding student teacher performance. The purpose of this study, then, was to investigate whether in fact an individual student teacher,

Claire Staab (PhD, Arizona State) is an Assistant Professor at the University of British Columbia, teaching classes in developmental reading which have a student teaching component. She has been involved in student teaching programs for several years both at UBC and Arizona State.

his/her university supervisor and his/her sponsor teacher tend to judge that student teacher's performance as being at the same level of competence. In order to carry out the investigation, it was necessary to construct a reliable measure of student teacher performance, which may in itself prove useful after further investigation.

It would seem that the literature should have an abundance of studies indicating how the student teaching experience is viewed by each member of the student teaching triad (i.e., the student teacher, his/her university supervisor, and his/her sponsor teacher), and how the opinions of each member of the triad compare with one another. On the contrary, the present researcher, as well as Hattie, Olphert, and Cole (1982), in an extensive review of the literature, found a paucity of research on this question. The few studies which have been done comparing the views of members of the triad have produced mixed results. Joyce (1967) found that student teachers and supervisory personnel did not agree on a student's competence in teaching behaviors except in the area of discipline, or class control. Buethe and Pettibone (1973), however, found a relatively high correlation of agreement between sponsor teachers and student teachers when the two groups were asked to judge specific aspects of the student teacher's performance. Perhaps this high correlation in the Buethe and Pettibone study (1973) resulted from the fact that only highly specific teaching behaviors such as spelling errors and the student teacher's dependence upon lecture notes were examined. Using only a small sample of thirteen student teachers, Simmons (1973) found a high correlation between the opinions of university supervisors and sponsor teachers regarding such student teaching behaviors as lesson planning and lesson execution. Student teachers were not asked to rate their own level of competence. Mahan and Lester (1974) found a relatively high correlation between the responses of student teachers, sponsor teachers and university supervisors when members of the triad were asked to make judgments regarding their satisfaction with the general effectiveness of the student teacher program in regard to such issues as amount of professional preparation and rapport with other members of the triad; however, no specific teaching behaviors were assessed.

Although there have been few studies comparing the views of members of the triad regarding the student teacher's performance, several studies have indicated specific concerns of student teachers, without comparing these concerns with those of supervisory personnel. Freeland (1979) has listed the three major concerns of student teachers as follows: motivating disinterested pupils; handling discipline problems; and evaluating pupil progress. Bartos and Lotven (1980), in asking student teachers to rank fifteen competencies in terms of those most needed in the classroom, listed the following in order of importance: the ability to apply appropriate techniques, the ability to maintain order and assist students in developing self-discipline; and the ability to motivate students. Purcell and Seifert (1982) found student teachers to be most concerned with the problem of student discipline. None of the above studies assessed the opinions of the sponsor teacher or the university supervisor regarding what problems they felt the student teacher was experiencing or what behaviors were most essential.

As previously mentioned, communication between members of the triad is necessary if desired changes are to be brought about in the student teacher's classroom performance. This communication is, in fact, vital to the student teaching experience (Yee, 1969; Seiferth and Solliday, 1980; Tabachnick, Popkewitz, & Zeichner, 1980; Queen & Mallen, 1982). This communication must be one in which there is an

understanding of each other's point of view and a frank and positive discussion of mutual concerns (Rickard, 1982).

It seems highly unlikely that members of the triad will understand each other's point of view if there is not some basic agreement about the student teacher's level of competence and some understanding of what each member of the triad views as a problem or a particular strength in the student teacher's behavior. For example, it would be difficult for a university supervisor to discuss a need for improved lesson planning with a student teacher if that student teacher viewed his/her performance in this behavior to be adequate. In order for a positive discussion to take place, student teachers and supervisory personnel must recognize how they as individuals view teaching behaviors and how these behaviors are viewed by other members of the triad. Even though several studies have established the fact that communication between members of the triad is necessary, adequate research regarding this first step of initial agreement is still lacking. Do student teachers and their supervisory personnel tend to view a student teacher's behavior at the same level of competence and thus have a common basis for beginning necessary dialogue?

In the present study sixty-six student teachers, ten university supervisors, and fifty-seven sponsor teachers were asked to respond to a questionnaire, the development of which is described below. This questionnaire sought to gain information on four specific clusters of teaching behaviors—Preparation for Instruction, Use of Teaching Skills, Classroom Management and Personal Qualities, in order to answer the following research questions:

1. Do student teachers judge various aspects of their teaching performance as being at the same level of competence as do their university supervisors and their sponsor teachers?
2. When relationships within the four clusters (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management and Personal Qualities) are compared, are the relationships among the three groups of subjects (i.e., student teachers, their university supervisors and their sponsor teachers) similar?
3. When relationships within single groups of subjects (e.g., student teachers) are examined, are the relationships among the four clusters (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management and Personal Qualities) similar?

Method

Subjects

Sixty-six student teachers, representing the entire population of a student teaching practicum, were asked to participate in the study. The practicum, comprising ten one-half day sessions in the classroom, was required for all elementary education students as an initial practicum and was conducted in the student teacher's second year of university study. Many of the student teachers had had previous experience in classrooms as observers, but this practicum was their first supervised teaching experience. The student teachers were assigned by the university student teaching office to an elementary school located within a large Western Canadian city or to a surrounding suburban area. The student teachers spent the ten sessions with one sponsor teacher in one classroom. The primary focus of this practicum was to prepare and teach lessons in reading, but the student teachers were also involved in helping the sponsor teacher with other subjects throughout the morning.

Fifty-seven sponsor teachers were asked to participate in the study. Some of these sponsor teachers supervised two student teachers simultaneously. Sponsor teachers were encouraged to conference with each student teacher weekly for at least one-half hour. These sponsor teachers comprised the total population of elementary school personnel supervising the student teachers.

Ten university supervisors participated in the study. These supervisors comprised the total population of university personnel supervising the student teachers. The university supervisors generally visited each student teacher in the classroom weekly and observed at least one entire lesson every two weeks. In most cases these university supervisors were also the student teacher's professor in a reading methods course which was taken simultaneously with the practicum. Student teachers also met with their university supervisors weekly for a one-hour seminar for the purpose of discussing problems encountered on the practicum, receiving feedback on lessons taught, and receiving additional help with teaching techniques. Individual conferences with the university supervisor were also scheduled as the student's need arose.

Procedure

At the end of the ten week practicum, student teachers, sponsor teachers and university supervisors were asked to respond to an identical questionnaire containing 16 items. This questionnaire is labeled Table 1. The 16 items were grouped into four clusters of teaching behaviors (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities). The respondents were asked to rate the student teacher's performance on these behaviors (i.e., separately the sponsor teacher and the university supervisor rated each of their own student teacher's performance and the student teacher rated his/her own performance). Each person responded to the questionnaire without any prior discussion with any other person. After data had been collected, each student teacher's responses were compared with those of his/her sponsor teacher and those of his/her university supervisor.

Neither the student teacher, the university supervisor, nor the sponsor teacher knew before the end of the practicum that they would be asked to fill out the questionnaire. To prevent collusion and to ensure that each questionnaire was answered independently by all respondents, certain precautions were taken. During the final week of the student teacher's experience the researcher visited the various student teaching seminars. The university supervisor conducting the seminar left the room and the student teachers responded to the questionnaires. These questionnaires were immediately collected by the researcher. Sixty-six student teachers were present in the seminars and each one responded to the questionnaire.

The following day the researcher asked each university supervisor to fill out the questionnaire on each student he/she supervised. These questionnaires were distributed and collected through university mail. All questionnaires distributed to university supervisors were returned.

The questionnaire was mailed to sponsor teachers during the week following the student teacher's last session. The sponsor teachers did not receive the questionnaire until contact with both the student teacher and the university supervisor had been terminated. A stamped, self-addressed envelope was provided for return to the researcher. All sponsor teachers except three returned the questionnaire.

All respondents were assured that their answers would not be available to anyone except the researcher and that their questionnaires would be coded to preserve anonymity.

TABLE 1
STUDENT TEACHER PERFORMANCE EVALUATION

UNIVERSITY SUPERVISOR & SPONSOR TEACHER FORM

Please fill in your name below proper title and the name of the student teacher whom you are evaluating.

Student Teacher	University Supervisor	Sponsor Teacher
Directions: Put an 'X' in the parentheses beside the number which corresponds to your evaluation of the student teacher's performance in each category. 5 is excellent. 1 is unsatisfactory.		

EXCELLENT	SATISFACTORY	UNSATISFACTORY
-----------	--------------	----------------

A. PREPARATION FOR INSTRUCTION

5 ()	4 ()	3 ()	2 ()	1 ()
-------	-------	-------	-------	-------

Lessons are consistently well prepared. Goal of lesson is thought through and teaching plans reflect this goal.	Lessons are adequately prepared. There is some goal to the lesson and teaching plans indicate some connection with goal.	Problems with lessons plans. Goals seem confused. Teaching plans are sometimes vaguely related.
---	--	---

5 ()	4 ()	3 ()	2 ()	1 ()
-------	-------	-------	-------	-------

Appropriate teaching materials are consistently present during lesson.	Appropriate materials are usually present during lesson.	Problems in having appropriate materials for lesson.
--	--	--

B. USE OF TEACHING SKILLS

5 ()	4 ()	3 ()	2 ()	1 ()
-------	-------	-------	-------	-------

Lesson plan is consistently executed with appropriate pace.	Lesson plan is usually executed with appropriate pace.	Problems in executing lesson plan with appropriate pace.
---	--	--

5 ()	4 ()	3 ()	2 ()	1 ()
-------	-------	-------	-------	-------

Lesson plan is consistently executed with appropriate sequence.	Lesson plan is usually executed with appropriate sequence.	Problems in executing lesson plan with appropriate sequence.
---	--	--

5 ()	4 ()	3 ()	2 ()	1 ()
-------	-------	-------	-------	-------

Instructional strategies are consistently varied to fit lesson.	Instructional strategies are usually varied to fit lesson.	Problems varying instructional strategies to fit lesson.
---	--	--

5 ()	4 ()	3 ()	2 ()	1 ()
-------	-------	-------	-------	-------

Examples and illustrations are consistently appropriate.	Examples and illustrations are usually appropriate.	Problems in using appropriate illustrations and examples.
--	---	---

Student Teachers' Perceptions of Competence

TABLE 1 (cont'd)

EXCELLENT		SATISFACTORY		UNSATISFACTORY
5 ()	4 ()	3 ()	2 ()	1 ()
Consistently reinforces learners in ways which increase their motivation.		Usually reinforces learners in ways which increase their motivation.		Problems reinforcing learners in ways which increase their motivation.
5 ()	4 ()	3 ()	2 ()	1 ()
Lesson is consistently aimed at the ability of the learners.		Lesson is usually aimed at the ability of the learners.		Problems aiming lesson at the ability of the learners.
C. CLASSROOM MANAGEMENT				
5 ()	4 ()	3 ()	2 ()	1 ()
Consistently anticipates difficulties and takes preventative measures.		Usually anticipates difficulties and takes preventative measures.		Problems anticipating difficulties and taking preventative measures.
5 ()	4 ()	3 ()	2 ()	1 ()
Consistently firm but fair in response to misbehavior.		Usually firm but fair in response to misbehavior.		Problems being firm but fair in response to misbehavior.
5 ()	4 ()	3 ()	2 ()	1 ()
Consistently helps learners to take responsibility for their own actions and practice self-discipline.		Usually helps learners take responsibility and practice self-discipline.		Problems in helping learners take responsibility for their own actions and practice self-discipline.
D. PERSONAL QUALITIES				
5 ()	4 ()	3 ()	2 ()	1 ()
Consistently enthusiastic about teaching and communicates this enthusiasm to learners.		Usually enthusiastic about teaching and communicates this enthusiasm to learners.		Problems in being enthusiastic about teaching and communicating this enthusiasm to learners.
5 ()	4 ()	3 ()	2 ()	1 ()
Consistently maintains good rapport with school staff.		Usually maintains good rapport with school staff.		Problems maintaining good rapport with school staff.
5 ()	4 ()	3 ()	2 ()	1 ()
Consistently willing to accept constructive criticism and institute changes.		Usually willing to accept constructive criticism and institute changes.		Problems receiving constructive criticism and instituting changes.

TABLE 1 (cont'd)

EXCELLENT		SATISFACTORY		UNSATISFACTORY
5 ()	4 ()	3 ()	2 ()	1 ()
Voice projection is consistently adequate for the circumstances of the lesson (e.g., size of group, size of room, etc.)		Voice projection is usually adequate for the circumstances of the lesson (e.g., size of group, size of room, etc.)		Problems with voice projection.
5 ()	4 ()	3 ()	2 ()	1 ()
General appearance is consistently adequate to meet the school's standards of professionalism.		General appearance is usually adequate to meet the school's standards of professionalism.		Problems in maintaining a general appearance which is adequate to meet the school's standards of professionalism.

Questionnaire Development

There were two slightly different formats of the questionnaire, one for the supervisors and sponsor teachers and one for the student teachers. The only difference in the formats was a change in verb form from third to first person (e.g., on the sponsor teacher and university supervisor's form, one item read, "Consistently helps students with self-discipline," while the student teacher's version read, "Consistently help students with self discipline." This slight change of verb form was necessary because student teachers were filling out the questionnaire in response to their own behavior and university supervisors and sponsor teachers were filling out the questionnaire in response to someone else's behavior (i.e., the student teacher's). The university supervisor and sponsor teacher form of the questionnaire appears as Table 1.

The 16 items on the questionnaire covered four main clusters of teaching performance and classroom behavior: Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities. These four clusters were obtained by reviewing the literature on necessary teaching skills (Peck & Tucker, 1973; Cooper, 1977; Chiarelott, Davidman, & Muse, 1980; Purcell & Seifert, 1982). Following this review of literature, twenty practicing elementary school teachers and five university professors involved in the supervision of student teachers were interviewed regarding the necessary teaching skills involved in student teaching. On the basis of this information and the literature review, a pilot questionnaire was prepared. This pilot questionnaire was given to eight university instructors and fifty-two student teachers. Neither these instructors nor these students were subjects in the final study. The respondents to the pilot questionnaire were asked to complete the instrument and also to indicate their opinion of its usefulness, its clarity of wording and format, and any problems in completing the form. The above procedures were necessary in establishing some estimate of content validity. The questionnaire which appears as Table 1 was developed as a result of this pilot questionnaire.

Analysis

A total of 50 questionnaires from each group (i.e., student teachers, university supervisors, and sponsor teachers) was analyzed. If either the student teacher, the

university supervisor, or the sponsor teacher did not complete the questionnaire or failed to complete all questions on the questionnaire, the data were not used.

Each of the four clusters which included Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities was treated separately. The results were analyzed as four clusters of teaching behaviors over three groups of subjects—student teachers, university supervisors, and sponsor teachers. An item analysis using Lertap (Nelson, 1974) was done to determine the Hoyt estimate of reliability, and each of the four clusters was subjected to a one-way analysis of variance across three groups of subjects. To determine the degree of similarity of opinion both between and within groups, Pearson correlations and tests of the significance of the difference between independently calculated correlations were calculated (Glass & Stanley, 1970).

TABLE 2
MEANS, STANDARD DEVIATIONS, AND
HOYT ESTIMATES OF RELIABILITY

Groups		Clusters of Teaching Behaviors			
		Preparation for Instruction	Use of Teaching Skills	Classroom Management	Personal Qualities
No.	of individual items in each cluster	2	6	3	5
1.	Students n=50	M=8.18 SD=(1.24) Hoyt=(0.67)	M=22.28 SD=(2.66) Hoyt=(0.60)	M=10.50 SD=(2.04) Hoyt=(0.73)	M=22.18 SD=(2.44) Hoyt=(0.66)
2.	University Supervisors n=50	M=7.86 SD=(2.13) Hoyt=(0.93)	M=22.30 SD=(5.08) Hoyt=(0.95)	M=10.62 SD=(2.65) Hoyt=(0.93)	M=21.08 SD=(4.07) Hoyt=(0.84)
3.	Sponsor Teachers n=50	M=8.78 SD=(1.50) Hoyt=(0.88)	M=22.88 SD=(3.26) Hoyt=(0.81)	M=10.30 SD=(2.22) Hoyt=(0.83)	M=21.56 SD=(3.39) Hoyt=(0.82)

Results

Internal Consistency of Questionnaire

The Hoyt estimates of internal consistency of the four clusters of teaching behaviors for each of the three groups of subjects (i.e., student teachers, sponsor teachers, and university supervisors) are presented in Table 2. The estimate for the sponsor teachers and the university supervisors are within or above the levels usually accepted (i.e., .70 or greater) for group decisions (Cronback, 1972). However, only those of the university supervisors are above or close to the suggested reliability coefficient for individual decision making (i.e., .95) (Cronback, 1972). The reliability coefficients for the four clusters were somewhat lower for the student group

(.60 to .73). This finding may be a result either of the brevity of the four cluster scales, or a reflection of the degree to which beginning student teachers perceive their teaching behaviors, as unified clusters.

Levels of Competence

Table 2 presents the means and standard deviations for the four clusters of teaching behaviors across the three groups of raters.

Table 3 presents the results of the analyses of variance for the four clusters of teaching behaviors over the three groups of subjects.

TABLE 3
ONE WAY ANALYSIS OF VARIANCE FOR FOUR CLUSTERS OF
TEACHING BEHAVIORS AS RATED BY
STUDENTS, UNIVERSITY SUPERVISORS AND SPONSOR TEACHERS

Sources of Variation	df	MS	F	p
1. Preparation for Instruction				
Between Groups	2	10.91	3.93	0.02
Within Groups	147	2.78		
2. Use of Teaching Skills				
Between Groups	2	5.80	0.40	0.67
Within Groups	147	14.50		
3. Classroom Management				
Between Groups	2	1.31	0.24	0.78
Within Groups	147	5.37		
4. Personal Qualities				
Between Groups	2	15.20	1.34	0.26
Within Groups	147	11.33		

The only significant difference between the three groups was achieved in Cluster One: Preparation for Instruction. Multiple comparisons were done using the Scheffe test at the .05 level of significance. On Cluster One the student teachers rated themselves similarly to their sponsor teachers. The sponsor teachers, however, rated the student teachers significantly higher than the university supervisors.

Inter-Cluster Relationships

Table 4 presents the between-group correlations for each of the four clusters of items for each possible pair of subjects (e.g., Student Teachers and Sponsor Teachers). Of the twelve correlations, nine achieved statistical significance beyond the .01 level. The remaining three were not significant. Within each item cluster, tests of the significance of the difference between correlations revealed that no pair was different for Cluster One: Preparation for Teaching, and Cluster Four: Personal Qualities. For Cluster Three: Classroom Management, agreement between university supervisors and sponsor teachers was higher than that between student teachers and sponsor teachers ($z=2.220, p<.05$), and that between student teachers and university supervisors ($z=2.839, p<.01$). Although statistical significance was

not achieved on any pair of correlations for Cluster Two: Use of Teaching Skills, the difference in correlation between student teachers and university supervisors, and university supervisors and sponsor teachers was marginal ($z = 1.53$, $p = 0.126$).

TABLE 4
BETWEEN GROUP PEARSON PRODUCT MOMENT CORRELATIONS OF
STUDENT TEACHERS, UNIVERSITY SUPERVISORS, AND SPONSOR
TEACHERS OVER FOUR CLUSTERS OF TEACHING BEHAVIORS
($n = 50$)

Groups	Clusters			
	1 Preparation for Instruction	2 Use of Teaching Skills	3 Classroom Management	4 Personal Qualities
Student Teachers to University Supervisors	0.52**	0.28	-0.03	0.45**
Student Teachers to Sponsor Teachers	0.53**	0.54**	0.16	0.45**
University Supervisors to Sponsor Teachers	0.41**	0.64**	0.58**	0.56**

* $p < .05$

** $p < .01$

Table 5 presents the within-group correlations across the four clusters for each of the three groups of subjects. As may be seen, all but one of the 18 coefficients achieved statistical significance ($p < .01$ in 16 cases and $p < .05$ in one case). None of the differences between pairs of correlations within each group of subjects was statistically significant, thereby lending support to the view that each group of subjects perceived student teaching behaviors as being relatively unified or “global.” For example, if a student teacher was perceived as being competent in one behavior, he or she was perceived similarly on other behaviors.

Discussion and Implication

In general, it appears that the self perceptions of student teachers regarding their teaching performance are similar in mean levels of competence to the perceptions of their university supervisors and sponsor teachers. There was some difference found across groups of subjects in only one of the four clusters: Preparation for Instruction (i.e., the teaching behaviors of preparing lessons and having appropriate teaching materials available). Sponsor teachers tended to rate student teachers more positively on Preparation for Instruction than did the student’s university supervisor. One possible explanation for sponsor teachers rating higher than university supervisors is that universities in their student teacher programs normally require extensive and complete lesson plans which are then checked by the university supervisor before the student actually enters the classroom. Often these lesson plans are far more extensive than those which have been developed by the

sponsor teacher. University supervisors tend to consider the preparation of lesson plans as a vital element in the training of student teachers and thus are often more critical of these plans. It may be necessary for the university supervisor to take more time at the beginning of each practicum to discuss mutual expectations for lesson plans with the sponsor teacher.

Examination of Table 4 reveals that 8 of the 12 correlations between groups of subjects on the four cluster scores were statistically significant, despite the fact that each correlation accounted for only a moderate percentage of the total variance (18% to 40%). Agreement was not consistent across the three groups of subjects for the clusters, Preparation for Instruction and Personal Qualities. Thus, although university supervisors perceived student teachers as being at different levels of competence on Cluster One: Preparation for Instruction, they did agree on relative ranking. This finding supports the conjecture made above about the higher expectations of university supervisors with regard to lesson planning.

Less agreement among the three pair of subjects was found on Cluster Two: Use of Teaching Skills and particularly on Cluster Three: Classroom Management. It is interesting that there tends to be less agreement on the teaching behaviors which a student teacher must demonstrate during the course of a lesson in progress (i.e., classroom management and such teaching skills as lesson execution, using appropriate examples, and adjusting the planned lesson to the learner's needs) than in those teaching behaviors which a student teacher may initially bring to the lesson (i.e., a prepared lesson plan and various personal qualities).

As the within-group correlations in Table 5 reveal, it was also found that university supervisors and sponsor teachers tend to view student teachers as performing consistently. If student teachers receive a high rating in one particular teaching behavior then they are likely to receive a high rating in other teaching behaviors. Students are not viewed as weak in lesson planning, for example, and strong in classroom management. The reason for this finding is speculative at this time. One possible explanation is that the four teaching behaviors, Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities may be highly interdependent. For example, if a student teacher's management skills are poor and the class is out of control it is highly likely that the lesson will not be well executed, appropriate materials will not be used, and personal qualities will suffer. Conversely, if a student has not planned a lesson adequately, does not implement the lesson with the proper sequence and pace, it is also likely that students will not understand what is being presented and tend to cause problems for the student teacher in the area of classroom management. Thus a deficiency in one teaching behavior may radiate to other teaching behaviors. If these teaching behaviors are interdependent, university supervisors and sponsor teachers must be trained to view a student teacher's performance and begin to pinpoint behaviors which are causing general problems. For example, if a student teacher's performance is deficient because of a poorly planned lesson, then immediate attention must be given to lesson planning, although perhaps there were classroom management problems during the execution of the lesson.

A second possible explanation for the finding that university supervisors and sponsor teachers tended to view student teachers as performing consistently is that students simply did perform similarly in different areas of teaching behaviors. Clearly, there are many students who have a "feel for teaching" and simply do demonstrate ability in all teaching behaviors, and conversely, there are those students who are clearly lacking in most teaching behaviors. A third possible explanation

would be the operation of “halo” effect, or a tendency to judge seemingly diverse behavior at a single level of competence. Further research is needed before the possible reason or reasons can be anything but speculative.

TABLE 5
WITHIN GROUP PEARSON PRODUCT MOMENT CORRELATIONS FOR
STUDENT TEACHERS, UNIVERSITY SUPERVISORS AND SPONSOR
TEACHERS OVER FOUR CLUSTERS OF TEACHING BEHAVIORS
(n=50)
Student Teachers

Clusters	Clusters			
	1 Preparation for Instruction	2 Use of Teaching Skills	3 Classroom Management	4 Personal Qualities
1		0.59**	0.22	0.58**
2			0.57**	0.62**
3				0.33*
4				

University Supervisors

Clusters	Clusters			
	1 Preparation for Instruction	2 Use of Teaching Skills	3 Classroom Management	4 Personal Qualities
1		0.92**	0.60**	0.81**
2			0.72**	0.84**
3				0.67**
4				

Sponsor Teachers

Clusters	Clusters			
	1 Preparation for Instruction	2 Use of Teaching Skills	3 Classroom Management	4 Personal Qualities
1		0.67**	0.60**	0.65**
2			0.71**	0.70**
3				0.67**
4				

* $p < .05$
** $p < .01$

To a lesser degree, student teachers also demonstrated global ratings in judging their own performance. As may be seen in Table 5, five of the six correlation coefficients for the student teachers achieved significance. However, the score for Cluster Three: Classroom Management was not significantly correlated with the student teachers' judgments of the remaining three clusters.

Possibly student teachers do tend to judge their performance in a relatively global manner, but single out the behavior of classroom management because it is of particular concern. This finding tends to support the findings in the Freeland (1979) and the Purcell and Seifert (1982) studies which noted that motivating disinterested pupils and handling discipline problems were student teachers' major concerns. The data from the present study do not indicate whether student teachers viewed classroom management as a major concern, but simply that they viewed their performance on this cluster differently from their performance on the other clusters of teaching behaviors. As a result of several informal conversations with students, it seems likely that classroom management was in fact a major concern. This speculation, combined with the results of the Freeland (1979) and Purcell and Seifert (1982) studies have definite implications for student teaching seminars. If classroom management is one of the student teacher's greatest concerns, both the university supervisor and the sponsor teacher must be sure to include ample information and instructions on this teaching behavior while conferring with students. Whether a generous amount of time is spent on the topic of classroom management is an issue for further research.

In the course of the study, a questionnaire of student teacher performance was constructed which may prove useful to other investigators, teacher trainers, and sponsor teachers. The instrument appears to be sufficiently reliable when used by university supervisors. However, further test development is necessary to improve the internal consistency of the questionnaire, when it is used by sponsor teachers or by student teachers to judge their own performance. In addition, although content validity was tentatively established through the methods reported above, estimates of empirical validity have yet to be measured. Following these steps, and the establishment of high inter-rater reliability coefficients, this questionnaire may be useful.

As this study was conducted at one institution and involved one student teacher training program, it is difficult to generalize the results to diverse populations of student teachers, university supervisors and sponsor teachers. Replication needs to be done in other institutions, in other training programs, and with student teachers who are not in their first practicum. As Borich and Madden (1977) have stated, the finding of no significant differences tends to be the norm in classroom research and more confidence can be placed in findings as they are replicated across studies using the same instrument. The study, however, has added to an extremely small body of research which has explored the question of whether members of the student teaching triad tend to judge the student teacher's performance as being at the same level of competence. In contrast to the Bueth and Pettibone (1973) study, this study used fairly inclusive measures of a student teacher's performance (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities); and in contrast to the Simmons (1973) study, this study used a larger sample (i.e., 50).

The study has also raised some interesting questions. How can university supervisors best communicate their expectations for lesson plans to sponsor teachers? Are most student teachers consistent in their degree of competence across

teaching behaviors or are they simply perceived by themselves and their supervisory personnel in such a manner? If in fact diverse competence exists, or if teaching behaviors are so interdependent that lack of competence in one behavior tends to radiate to lack of competence in other behaviors, how can supervisory personnel be trained to perceive specific problems and communicate them to student teachers? Do most student teachers perceive classroom management as their greatest weakness? If so, how much help is being provided by supervisory personnel in improving this teaching behavior? In general, it appears that student teachers, university supervisors, and sponsor teachers do tend to judge the student teacher's performance as being at the same level of competence, yet answers to the above questions would certainly help to increase the necessary communication within the student teaching triad.

References

- Bartos, R., & Lotven, B. (1980). *Need and proficiency of teacher competencies as perceived by student teachers vs. practicing professional teachers*. (ERIC Document Reproduction Service No. ED 193 179)
- Borich, G., & Madden, S. (1977). *Evaluating classroom instruction: A sourcebook of instruments*. Reading, MA: Addison-Wesley.
- Bueth, C., & Pettibone, T. (1973). Evaluating student teaching behavior: Three views. *Teaching Education*, 9, 121-125.
- Campbell, L., & Williamson, J. (1973). Practical problems in the student teacher cooperative teacher relationship. *Education*, 94, 168-169.
- Chiarelott, L., Davidman, L., & Muse, C. (1980). Evaluating preservice teacher candidates. *The Clearing House*, 53, 295-299.
- Cooper, J. (Ed.). (1977). *Classroom teaching skills: A handbook*. New York: Harper & Row.
- Cronback, L. J. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- Freeland, K. (1979). Assessing student teaching. *Teacher Education*, 15, 11-16.
- Glass, G. V., & Stanley, J. C. (1970). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Hattie, J., Olphert, W., & Cole, B. (1982). Assessment of student teachers by supervising teachers. *Journal of Educational Psychology*, 74, 778-785.
- Joyce, D. (1967). *Exploration of the utilization of personnel in the supervision of student teachers when feedback via films and systems for the analysis of teaching are introduced into the student teaching program* (Final Report). (ERIC Document Reproduction Service No. ED 020 575)
- Larson, E., & Ringness, T. (1965). *Effect of clinical evaluation on supervision of student teachers*. (ERIC Document Reproduction Service No. ED 003 467)
- Mahan, J., & Lester, F. (1974). *A study of inter-relationships between three independent rankings of student teacher effectiveness and student teacher assessments of satisfaction*. Chicago: Annual meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED 096 323)
- Nelson, L. (1974). *Guide to Letrap use and interpretation*. Dunedin, New Zealand: Department of Education, University of Otago.
- Peck, R., & Tucker, J. (1973). Research in teacher education. In R. Travers (Ed.), *The second handbook of research on teaching*. Chicago: Rand McNally.
- Purcell, T., & Seifert, B. (1982). A tri-state survey of student teachers. *College Student Journal*, 16, 27-29.

- Queen, J., & Mallen, L. (1982). Student teaching: A transitional approach. *The Clearing House*, 56, 56-59.
- Reynolds, J. (1977). A modified approach to teacher education assessment. *Teacher Education*, 13, 9-13.
- Rikard, G. (1982). The student teacher practicum. *Journal of Physical Education, Recreation and Dance*, 53, 60-61.
- Seiferth, B. M., & Solliday, M. A. (1980). A time study of the patterns of activities of student teachers. *College Student Journal*, 14, 285-287.
- Simmons, J. (1973). Two approaches: Name your poison. *English Education*, 4, 151-160.
- Tabachnik, B., Popkewitz, T., & Zeichner, K. (1980). Teacher education and the professional perspectives of student teachers. *Interchange*, 10, 12-19.
- Yee, A. (1969). Do cooperating teachers influence the attitudes of student teachers? *Journal of Educational Psychology*, 60, 328-332.

HORST H. MUELLER

University of Alberta

GABRIEL J. MANCINI

Edmonton Public School Board

and

ROBERT H. SHORT

University of Alberta

An Evaluation of the Diagnostic Efficiency of the Wechsler Intelligence Scale for Children—Revised

Although not intended as such, the WISC-R has been used extensively as a diagnostic instrument even though its effectiveness in such a role has remained in doubt. This study assessed the diagnostic suitability of the WISC-R by applying Kelley's (1923) method for judging the orthogonality of tests through the estimation of the proportion of score differences in excess of chance to all WISC-R subtest paired comparisons. Kelley's method was also applied to factor comparisons, Bannatyne (1974) cluster comparisons, and WISC-R Scale comparisons at each of the 11 age levels for which the WISC-R was standardized.

The results indicate that caution needs to be used when applying the WISC-R diagnostically. More than half of all possible subtest comparisons proved inappropriate for diagnostic purposes while others varied according to the age level in question. The factor, cluster, and Scale comparisons proved somewhat better for diagnosis. Generally, the findings of this study support the contention that three groupings of WISC-R subtests based on Kaufman's (1975) factor analysis are fully adequate for describing WISC-R performance. The authors advocate the use of the comparison tables presented in this paper as a necessary preliminary in any attempt to use the WISC-R as a diagnostic tool.

Horst Mueller (MA, University of Saskatchewan) is presently a doctoral student in the Department of Educational Psychology at the University of Alberta. His research interests include behavior therapy with children, self-management training and behavioral medicine, clinical assessment, and reading comprehension processes.

Gabriel Mancini is a consulting psychologist with the Edmonton Public School Board, Edmonton, Alberta. His research interests are metacognition and behavioral disorder in children.

Robert Short is Associate Professor in the Department of Educational Psychology at the University of Alberta. His research interests include human learning and development, individual differences, and instructional psychology.

The Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1974) is currently the most popular and well-researched psychometric instrument for the assessment of the intellectual functioning of individual children in both clinical and educational settings (Kaufman, 1980; Quattrocchi & Sherrets, 1980; Sattler, 1982). This popularity has come about in part because, although designed solely as a measure of "global intelligence" or an individual's "overall capacity for intelligent behavior" (Wechsler, 1974, p. 1), the WISC-R also has a number of features which render it attractive for use as a psychodiagnostic tool. Such features as the excellent reliability of its three Scale scores; the "adequate" reliable specific variance of most of its subtests (Dinning, 1976; Kaufman, 1980); the relative stability and remarkable similarity of its factor composition across age (Kaufman, 1975), sex, (Reynolds & Gutkin, 1980), intelligence levels (Karnes & Brown, 1980; Van Hagen & Kaufman, 1975), ethnic/racial groupings (Reschley, 1978), and clinical diagnostic groupings (Kaufman, 1979, 1981; Lutey, 1977; Quattrocchi & Sherrets, 1980; Sattler, 1982); together with its division into separate verbal and nonverbal scales have all greatly added to the WISC-R's clinical popularity.

In recent years, more and more research has focused on the WISC-R's potential as an educational and clinical differential diagnostic tool, and many clinicians have come to view the pattern (high versus low scores) of an individual's performance on the WISC-R subtests as potentially indicative of specific mental operations or brain-behavior relationships (Kaufman, 1980; Sattler, 1982). In this regard a considerable literature has accumulated on various WISC-R test scatter and subtest profile indices and their association with specific exceptionalities (e.g., Anderson, Kaufman, & Kaufman, 1976; Bannatyne, 1968, 1974; Clarizio & Bernard, 1981; Dash, Dennis, Mueller, Mancini, Snart, & Short, 1983; Dean, 1977, 1978; Groff & Hubble, 1981; Hale, 1979; Hale & Landino, 1981; Haynes & Bensch, 1981; Kunc & McMahon, 1979; Loro & Woodward, 1976; McIntosh, 1974; Mueller, Matheson, & Short, 1983; Tew, 1977; Thompson, 1980; Thompson, 1981; Vance & Singer, 1979; Webster & Lafayette, 1980). Unfortunately, the results of such studies have been mixed and have failed to provide any firm basis for making diagnostic classification decisions on the basis of WISC-R subtest scatter or performance profiles (Kaufman, 1981; Lutey, 1977; Sattler, 1982). A number of authors have also leveled criticism at both the usefulness of such research and at some of the basic assumptions underlying the use of scatter and profile analysis approaches to differential diagnosis and classification of children with the WISC-R (e.g., Gutkin, 1979a; Hirshoren & Kavale, 1976; Miller, 1980; Ryckman, 1981). A major caution stated by many authors is that the reliabilities of the individual WISC-R subtests are too low to warrant interpretation of performance profiles based on individual subtest scores (Hirshoren & Kavale, 1976). However, the practice of WISC-R profile analysis, based on either individual subtests or clusters of subtests, continues among both researchers and clinicians (Kaufman, 1979; Sattler, 1982).

The problem of diagnosis may be seen as twofold. Diagnosis is both the process of classifying individuals on the basis of their problems or difficulties, and the process of determining the nature of the problem or difficulty so as to guide appropriate intervention or remediation (see Chaplin, 1975). Similarly, a diagnostic test is any instrument which is utilized in an effort to classify individuals on the basis of the difficulties they demonstrate and/or to determine the nature or source of such individuals' problems (Salvia & Ysseldyke, 1981). The WISC-R has been used both in attempts to classify individuals on the basis of their subtest performance profiles (e.g., Vance, Wallbrown, & Blaha, 1978; Vance & Singer, 1979), and to determine

specific areas of difficulty and suggest remediation strategy for children already classified (e.g., Wallbrown, Vance, & Blaha, 1979). The first of these usages involves inter-individual comparisons of subtest performance profiles, whereas the second usage involves intra-individual comparisons. Thus the WISC-R must be capable of determining "true" differences in performance profiles of individuals from the norm and, at the same time, it must be able to determine "true" or reliable intra-individual differences in abilities. In a sense, any normative-based psychodiagnostic instrument must be adequate for both these uses.

A major failing within this research area has been the lack of studies examining whether or not the WISC-R, which was designed to reliably differentiate individuals on the basis of its three Scale scores, is indeed capable of differentiating individuals on the basis of differential scores across subtest pairs or across clusters of subtests based on rational (Bannatyne, 1968, 1974) or factorial (Kaufman, 1975) grouping schema. This is an important question because, underlying the profile analysis research is the often less-than-explicit notion that a specific pattern of differential WISC-R performance that is reliable for the individual will also reliably distinguish that individual from the normative group. The usual approach to using the WISC-R diagnostically has been to simply examine the individual's profile of subtest or subtest-cluster scores for score differences large enough to be significant at the .05 or .01 level of probability as determined from various published tables of significant score discrepancies (Gutkin, 1978, 1979b; Kaufman, 1979; Lutey, 1977; Piotrowski & Grubb, 1976; Reynolds, 1981; Sattler, 1982). The application of such procedures to the test results of an individual assumes that the subtests or subtest-clusters have been found, by an analysis of results for the population, capable of reliably showing such differences. Such an analysis has yet to be done for the WISC-R.

For any test or test-battery to demonstrate adequate diagnostic efficiency it must be comprised of subtests which are not highly intercorrelated and which are individually quite reliable. Such a combination of psychometric properties will ensure that the test is capable of measuring relatively independent traits and that any obtained score differences between subtests will be attributable to true individual differences in the traits being measured rather than to measurement error (Salvia & Ysseldyke, 1981). The WISC-R seems a rather dubious candidate for an efficient diagnostic test because it is composed of subtests that have a relatively large amount of common variance with less individual specific variance (Dinning, 1976; Sattler, 1982), and its individual subtests vary greatly in reliability (Wechsler, 1974).

The present paper is concerned specifically with an analysis and evaluation of the differential diagnostic efficiency of the WISC-R and will apply Kelley's (1923) method for estimating the proportion of differences in excess of chance to all possible WISC-R subtest paired comparisons, Scale comparisons, Kaufman (1975) factor comparisons, and Bannatyne (1974) cluster comparisons at each of the 11 age levels for which the WISC-R was standardized.

Although other procedures for assessing the reliability of score differences between tests are available (Mollenkopf, 1950; Mosier, 1951; Nunnally, 1967), Kelley's method for determining the proportion of differences in excess of chance has the advantage of being directly meaningful and easily understandable (French, 1966). Unfortunately, apart from its use in studies by Segel (1934), Bennett and Doppelt (1948), and Tierney and Ames (1978), this particular technique has been rarely ap-

plied to test data. Based on the reliabilities of the tests and their intercorrelations, the index derived by Kelley's method is equivalent to the percentage of time differences between two tests represent true differences in ability.

A better understanding of Kelley's approach for estimating the diagnostic efficiency of two tests or batteries of tests may be gained by supposing that two persons stand at exactly the same level on some aptitude. When these two people are tested for this aptitude by fallible tests, there will be a difference in the scores they receive. If such testing is done repeatedly, a distribution of difference scores will be obtained. This distribution, of course, will be entirely attributable to chance. In cases where real differences in aptitude exist between two people, the observed differences in scores will be greater and will be partly attributable to chance and partly attributable to real differences in aptitude. The efficiency of a diagnostic test can be stated in terms of the proportion of observed differences that are not attributable to chance. If the two variables in question are highly related, the real differences will be rather small and, therefore, the proportion of differences not accounted for by chance will be low. If, on the other hand, the two variables are highly independent, the differences will be larger. Test reliability also comes into play here in that if the two tests have poor reliability, a greater proportion of differences will be due to chance. Thus, as stated earlier, a good differential diagnostic test must be comprised of subtests that are highly independent of each other and highly reliable. Kelley's (1923) formula permits a determination of the diagnostic efficiency of a test through the determination of that proportion of differences between two subtests or batteries of tests, based on the tests' intercorrelation and reliabilities, that is in excess of chance. For example, for two tests a proportion of differences in excess of chance of .35 indicates that fully 65 percent of all score differences would be due to measurement error. The larger the proportion of differences in excess of chance, the more diagnostically efficient the test.

Segel (1934), in his research with Kelley's procedure, concluded that tests exhibiting a proportion of score differences in excess of chance less than .25 should not be considered diagnostic in relation to each other—such tests having little value for differential diagnosis. Segel's minimum criterion for diagnostic efficiency may at first glance appear overly lenient, since at the .25 level he advocates fully 75 percent of all score differences between two tests would be due to chance. However, a review of the few studies which have examined the diagnostic efficiency of tests using the Kelley method (Bennett & Doppelt, 1948; Bennett, Seashore, & Wesman, 1947; Kelley, 1923; Segel, 1934; Tierney & Ames, 1978) reveals that few tests are capable of differentiating abilities at much better than 25 percent above chance. For example, Kelley (1923) discovered that the Stanford Achievement Test was capable of differentiating arithmetic (total arithmetic score) from reading (total reading score) abilities at no more than 36 percent above chance. Similarly, Segel (1934) found that comparisons between the three parts of the Gates Primary Reading Test and the four parts of the Gates Silent Reading Test demonstrated a mean diagnostic efficiency of barely more than 25 percent above chance. More recently, Tierney and Ames (1978), using the Kelley procedure to evaluate the diagnostic efficiency of the Illinois Test of Psycholinguistic Abilities, reported "that 45 out of the 66 possible subtest comparisons were . . . inappropriate [below .25] at at least one age level" (p. 589). Such research results clearly show the Segel criterion for diagnostic efficiency to be not overly liberal.

Method

The Kelley (1923) procedure for estimating the proportion of score differences between two tests in excess of chance was applied to the WISC-R standardization data (Wechsler, 1974). Using Segel's (1934) criterion of 25 percent of score differences in excess of chance as a minimum standard for diagnostic utility, the diagnostic efficiency of WISC-R subtest paired comparisons, WISC-R Scale comparisons, Kaufman (1975) factor grouping comparisons, and Bannatyne (1974) cluster comparisons were evaluated for each of the 11 age groups of the standardization population.

First, using the subtest reliabilities and intercorrelation coefficients given in the WISC-R manual (pp. 28, 36-47), Kelley's formula for estimating tests' diagnostic efficiency was applied to each of the 66 unique WISC-R subtest paired comparisons at each of 11 age levels and the proportion of differences in excess of chance was calculated for each of these 726 comparisons. A similar set of calculations was carried out for the WISC-R Verbal Scale versus Performance Scale comparisons at each of the 11 age levels using the Scale reliabilities and intercorrelation coefficients published in the WISC-R manual.

For the evaluation of the diagnostic efficiency of the three WISC-R factor groupings of subtests first derived by Kaufman (1975), a more complex procedure was followed. First, a composite reliability was estimated for each of the three factor groupings of subtests at each of 11 age levels using the general formula developed by Mosier (Guilford, 1954, p. 393). The intercorrelation of the three factor groupings of subtests at all 11 age levels was also estimated by transforming the composite subtest intercorrelation coefficients given in the WISC-R manual to Fisher *Z*-scores (Smith, 1970), obtaining their mean for each factor grouping, and then converting this mean *Z*-score back into an intercorrelation coefficient. These composite reliabilities and intercorrelation estimates for the three factor groupings of subtests were then used in Kelley's formula to calculate the proportion of score differences in excess of chance for all possible paired comparisons of the Verbal Comprehension (Information, Similarities, Vocabulary, Comprehension), Perceptual Organization (Picture Completion, Picture Arrangement, Block Design, Object Assembly), and Freedom from Distractibility (Arithmetic, Digit Span, Coding) factors at each of 11 age levels.

Lastly, a similar procedure to that used to estimate the diagnostic efficiency of Kaufman's WISC-R factors was used to estimate the diagnostic efficiency of the four Bannatyne (1974) WISC-R subtest clusters. As before, the proportion of score differences in excess of chance was calculated for all possible paired comparisons of the Bannatyne clusters at each of 11 age levels.

Results

Table 1 presents the proportion, expressed as a percentage, of score differences in excess of chance for WISC-R subtest paired-comparisons at each of the 11 age levels for which the WISC-R was standardized. From this table it can be seen that each of the individual WISC-R subtests are involved in at least 38 inappropriate comparisons. At certain age levels, some of the subtests (e.g., Information at the 6.5 year level) are never involved in an appropriate comparison. The subtests which fare best are Digit Span, Coding, and Block Design where 69 percent, 58 percent, and 50 percent of comparisons, respectively, meet or exceed Segel's .25 criterion for minimal diagnostic efficiency. On the other hand, Object Assembly, Comprehen-

TABLE 1
PERCENTAGE OF SCORE DIFFERENCES IN EXCESS OF CHANCE
FOR ALL WISC-R SUBTEST PAIRED-COMPARISONS

Subtest Pairs	Age Level										
	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5	14.5	15.5	16.5
Info.-Simil.	21	22	17	12	18	16	21	16	17	17	24
Info.-Arith.	11	13	19	19	21	24	26*	24	16	30*	21
Info.-Voc.	11	14	21	12	19	17	19	17	19	23	22
Info.-Comp.	15	17	18	18	18	23	26*	21	23	22	22
Info.-D.Sp.	15	27*	34*	22	26*	29*	29*	31*	31*	35*	30*
Info.-P.Com.	23	28*	32*	26*	23	29*	23	28*	23	25*	27*
Info.-P.Arr.	14	19	18	23	26*	25*	25*	23	25*	32*	28*
Info.-B.Des.	17	23	29*	21	31*	32*	28*	29*	30*	36*	37*
Info.-O.Ass.	19	18	21	21	19	24	19	30*	24	28*	25*
Info.-Cod.	17	23	21	28*	32*	31*	34*	24	27*	37*	32*
Info.-Maz.	21	29*	31*	27*	26*	30*	22	32*	31*	28*	24
Simil.-Arith.	29*	25*	21	22	19	22	26*	18	17	21	23
Simil.-Voc.	23	14	21	10	11	14	19	12	15	14	22
Simil.-Comp.	23	21	15	13	13	17	21	14	13	11	17
Simil.-D.Sp.	27*	33*	32*	18	24	30*	28*	27*	25*	24	29*
Simil.-P.Com.	33*	31*	30*	22	17	24	21	20	17	15	20
Simil.-P.Arr.	28*	23	21	18	22	24	26*	20	20	19	25*
Simil.-B.Des.	32*	27*	27*	16	24	26*	26*	25*	23	23	31*
Simil.-O.Ass.	30*	23	21	14	15	19	17	24	19	19	24
Simil.-Cod.	29*	27*	14	27*	26*	27*	32*	21	23	26*	28*
Simil.-Maz.	33*	35*	24	23	24	26*	21	27*	25*	19	21
Arith.-Voc.	17	14	25*	21	23	26*	24	25*	22	28*	25*
Arith.-Comp.	18	17	21	22	21	25*	27*	22	20	21	21
Arith.-D.Sp.	19	23	23	19	21	24	19	24	21	25*	17
Arith.-P.Com.	29*	29*	27*	27*	22	27*	24	22	21	22	25*
Arith.-P.Arr.	23	21	19	27*	24	26*	27*	27*	22	28*	25*
Arith.-B.Des.	23	24	24	21	28*	28*	22	27*	23	28*	30*
Arith.-O.Ass.	23	20	19	24	21	24	19	28*	23	27*	26*
Arith.-Cod.	21	22	14	25*	25*	30*	27*	22	21	30*	25*
Arith.-Maz.	28*	28*	24	25*	21	25*	18	26*	26*	24	19
Voc.-Comp.	12	06	15	14	13	16	24	16	17	14	19
Voc.-D.Sp.	23	24	40*	22	28*	29*	26*	30*	30*	30*	30*
Voc. P.Com.	23	23	31*	28*	19	27*	24	28*	26*	25*	27*
Voc.-P.Arr.	18	17	20	23	24	26*	24	25*	26*	30*	28*
Voc.-B.Des.	21	21	34*	24	28*	32*	25*	31*	33*	36*	39*
Voc.-O.Ass.	21	20	25*	20	17	24	18	29*	28*	30*	28*
Voc.-Cod.	20	18	24	28*	30*	32*	32*	24	28*	34*	34*
Voc.-Maz.	27*	26*	33*	28*	27*	31*	21	32*	32*	28*	24*

sion, and Similarities fare least well with only 20 percent, 28 percent, and 31 percent of comparisons being diagnostically tenable. A total of 11 of the 66 possible subtest paired-comparisons are diagnostically valueless at all age levels. These are:

TABLE 1 (CONT'D)

Subtest Pairs	Age Level										
	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5	14.5	15.5	16.5
Comp.-D.Sp.	21	25*	33*	22	26*	29*	31*	28*	29*	30*	29*
Comp.-P.Com.	24	23	27*	19	17	26*	23*	19	16	16	21
Comp.-P.Arr.	18	17	13	19	21	24	28*	23	21	21	24
Comp.-B.Des.	19	21	27*	20	25*	32*	31*	28*	27*	24	30*
Comp.-O.Ass.	21	17	18	15	18	23	23	25*	23	20	22
Comp.-Cod.	21	21	19	28*	26*	30*	32*	23	24	26*	26*
Comp.-Maz.	21	23	23	26*	21	28*	24	27*	28*	18	17
D.Sp.-P.Comp.	30*	36*	34*	25*	27*	32*	24	32*	28*	27*	29*
D.Sp.-P.Arr.	23	26*	32*	27*	25*	27*	30*	27*	24	29*	29*
D.Sp.-B.Des.	25*	32*	37*	24	26*	34*	26*	32*	33*	32*	35*
D.Sp.-O.Ass.	25*	28*	28*	23	21	27*	22	31*	26*	27*	28*
D.Sp.-Cod.	21	27*	22	23	22	24	27*	19	24	30*	30*
D.Sp.-Maz.	28*	32*	34*	25*	22	31*	21	25*	27*	24	22
P.Com.-P.Arr.	23	24	24	19	21	23	16	21	21	17	22
P.Com.-B.Des.	24	27*	32*	18	19	27*	18	18	16	17	22
P.Com.-O.Ass.	24	21	18	14	10	16	12	18	14	16	15
P.Com.-Cod.	23	27*	23	29*	28*	35*	30*	24	23	28*	30*
P.Com.-Maz.	32*	26*	31*	21	17	27*	15	21	19	16	18
P.Arr.-B.Des.	18	18	21	21	26*	28*	25*	21	22	24	26*
P.Arr.-O.Ass.	18	14	10	16	16	18	14	19	21	19	18
P.Arr.-Cod.	19	18	17	27*	26*	26*	29*	23	21	30*	28*
P.Arr.-Maz.	24	24	21	23	20	24	17	23	21	19	17
B.Des.-O.Ass.	18	11	13	10	12	13	13	16	13	13	14
B.Des.-Cod.	19	23	21	28*	30*	33*	29*	23	26*	30*	34*
B.Des.-Maz.	24	23	28*	18	19	27*	20	21	24	19	16
O.Ass.-Cod.	19	23	17	24	24	28*	23	21	22	26*	26*
O.Ass.-Maz.	18	21	17	17	19	24	11	17	22	16	14
Cod.-Maz.	27*	23	23	26*	24	32*	24	19	19	25*	22

Note:Asterisks denote those comparisons with a diagnostic efficiency of .25 or better.

Information with either Similarities or Vocabulary, Similarities with either Vocabulary or Comprehension, Vocabulary with Comprehension, Picture Completion with either Picture Arrangement or Object Assembly, Picture Arrangement with either Object Assembly or Mazes, and Object Assembly with either Block Design or Mazes. Generally, score differences between individual subtest-pairs within either the Verbal Comprehension factor or the Perceptual Organization factor (Kaufman, 1975) are most likely to be due to chance. All of the 66 possible subtest paired-comparisons failed to meet Segel's minimal requirement for diagnostic efficiency at at least one age level. Clearly, basing diagnostic statements upon score differences between individual WISC-R subtests is a rather risky proposition.

TABLE 2
PERCENTAGE OF SCORE DIFFERENCES IN EXCESS OF CHANCE FOR
WISC-R SCALE COMPARISONS, KAUFMAN FACTOR
COMPARISONS, AND BANNATYNE CLUSTER COMPARAISONS

	Age Level										
WISC-R Scales	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5	14.5	15.5	16.5
Verb.-Perf.	32*	32*	33*	31*	31*	34*	33*	36*	32*	37*	36*
Kaufman Factors											
VC-PO	39*	43*	41*	37*	41*	45*	45*	45*	44*	41*	47*
VC-FD	41*	42*	47*	45*	41*	45*	51*	45*	45*	47*	53*
PO-FD	42*	45*	42*	45*	36*	47*	47*	44*	45*	45*	51*
Bannatyne Clusters											
Con.-Spat.	40*	37*	42*	37*	36*	43*	39*	42*	42*	39*	41*
Con.-Seq.	36*	36*	40*	39*	39*	42*	42*	39*	37*	41*	42*
Con.-Ac.Kn.	16	08	14	12	13	08	13	14	18	16	16
Spat.-Seq.	37*	39*	37*	39*	38*	43*	39*	40*	37*	41*	43*
Spat.-Ac.Kn.	37*	37*	40*	41*	39*	43*	39*	44*	42*	44*	43*
Seq.-Ac.Kn.	09	08	15	14	16	19	14	17	13	22	16

Note:

1. Asterisks denote comparisons with a diagnostic efficiency of .25 or better.
2. VC = Verbal Comprehension, PO = Perceptual Organization, FD = Freedom from Distractibility, Con. = Conceptual, Spat. = Spatial, Seq = Sequential, and Ac.Kn. = Acquired Knowledge

An examination of the diagnostic efficiency of various groupings or clusterings of WISC-R subtests presents a somewhat brighter outlook than that for individual subtest pairings. Table 2 presents the percentage of score differences in excess of chance for WISC-R Scale comparisons, factor comparisons, and Bannatyne cluster comparisons at each of 11 age levels. From this table it can be seen that comparisons between the WISC-R Verbal Scale and Performance Scale are diagnostically tenable at all age levels with an efficiency of at least 30 percent above chance. This apparent diagnostic efficiency of the WISC-R's two Scales is even further improved when the WISC-R is divided into three scales based on Kaufman's (1975) factor analytic research with the WISC-R. The use of a three-scale interpretation permits the differentiation of abilities at better than 40 percent above chance for most of the possible comparisons.

An examination of Table 2 also reveals that for the four Bannatyne clusters of subtests only certain paired comparisons are tenable at all age levels, whereas others are not. The three Bannatyne clusters that most closely resemble the Kaufman WISC-R factor groupings, that is, the Conceptual, Spatial, and Sequential clusters, all demonstrate adequate diagnostic efficiency at all age levels when paired with each other. On the other hand, the fourth Bannatyne cluster, Acquired Knowledge, lacks diagnostic value when involved in comparisons with either

the Conceptual or Sequential clusters. This lack of diagnostic efficiency for such pairings is a function of the fact that Acquired Knowledge contains subtests that are also found in the Conceptual and Sequential clusters, thus dramatically increasing the intercorrelation of Acquired Knowledge with both Conceptual and Sequential (Matheson, Mueller, & Short, in press). Table 2 also reveals that, generally, paired comparisons between Bannatyne clusters of subtests are less diagnostically efficient than paired comparisons between Kaufman factor groupings at most age levels. This is due to the somewhat lower composite reliability of the Bannatyne clusters.

Discussion

The analysis of WISC-R standardization data (Wechsler, 1974) using Kelley's (1923) procedure for the determination of the proportion of all score differences between tests in excess of chance permitted an evaluation of the diagnostic efficiency of the WISC-R subtests and Scales, Kaufman's (1975) WISC-R factor groupings, and Bannatyne's (1974) recategorized clusters of WISC-R subtests. The application of Segel's (1934) criterion for minimal diagnostic efficiency to the results of the data generated by Kelley's procedure served to reference those subtest, factor group, and cluster comparisons that may be considered diagnostically tenable at the age levels for which the WISC-R was originally standardized. Assuming the logic of our analyses and procedures, reference to Tables 1 and 2 should precede any attempt to use the WISC-R in a diagnostic manner. If, according to the data presented in our tables, certain subtest or cluster comparisons are diagnostically valuable, then appropriate standard error of measurement confidence intervals (see Gutkin, 1978; Reynolds, 1981; Sattler, 1982) may be applied to determine if a score difference is statistically significant. However, while the acceptance of the value of a particular subtest or cluster comparison does not negate taking standard differential diagnostic precautions (see Sattler, 1982), the rejection of a particular comparison on grounds of poor diagnostic efficiency does suggest that the compared tests cannot accurately discriminate real inter-individual differences in measured abilities from score differences caused by measurement error, regardless of the size of such differences. The chances of statistically significant differences being caused by measurement error are far too great for test comparisons with a diagnostic efficiency of less than 25 percent to have any value in differential diagnosis. For these reasons, reference to the tables presented herein should be a first step in any attempt to use the WISC-R as a diagnostic tool.

More generally, the results of the present analysis lend further support to previously expressed cautions against differential diagnosis based on WISC-R subtest comparisons (e.g., Hirshoren & Kavale, 1976; Lutey, 1977; Sattler, 1982; Wechsler, 1974), and to those who argue that three groupings of subtests are fully adequate for describing performance on the WISC-R (Hirshoren & Kavale, 1976; Kaufman, 1975). While the Bannatyne clusters appear to permit four tenable comparisons compared to only three using the Kaufman factor groupings, the Kaufman groupings show the highest diagnostic efficiency at all age levels. Certainly the four-cluster Bannatyne (1974) recategorization of the WISC-R subtests would appear to add little to the Kaufman three-factor solution other than permitting a tenable comparison between the Spatial and Acquired Knowledge clusters. This finding was given further confirmation recently when multiple group factor analysis was used to test Bannatyne's (1974) categories against the subtest correlations for the 7.5, 10.5,

and 13.5 year samples from the WISC-R standardization. In Matheson, Mueller, and Short's (in press) study, it was concluded "that Bannatyne's Acquired Knowledge category has no validity as a Construct which is separate from his conceptualization category" (p. 22).

In conclusion, it seems clear that although the WISC-R was not designed as a diagnostic test, it does show some efficiency in differentiating children on the basis of at least three broad abilities or aptitudes. However, in most cases, diagnostic attempts based on score differences between individual WISC-R subtests must be viewed as extremely poor practice.

Notes

1. Reprints may be obtained from the first author at the Department of Educational Psychology, Faculty of Education, University of Alberta, Edmonton, Alberta, Canada.
2. This paper was partially funded by a Social Sciences and Humanities Research Council of Canada Doctoral Fellowship awarded to the first author.
3. The authors wish to acknowledge the assistance of Udaya Dash, Steven Dennis, and Dwight Harley for their help in data analysis and programing, and editing.

References

- Anderson, M., Kaufman, A., & Kaufman, N. (1976). Use of the WISC-R with a learning disabled population: Some diagnostic implications. *Psychology in the Schools*, 13(3), 381-386.
- Bannatyne, A. (1968). Diagnosing learning disabilities and writing remedial prescriptions. *Journal of Learning Disabilities*, 1(4), 28-35.
- Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC-R scaled scores. *Journal of Learning Disabilities*, 7(2), 272-274.
- Bennett, G., & Doppelt, J. (1948). The evaluation of pairs of tests for guidance use. *Educational & Psychological Measurement*, 8(3), 319-325.
- Bennett, G., Seashore, H., & Wesman, J. (1947). *Differential aptitude tests*. New York, NY: Psychological Corporation.
- Chaplin, J. (1975). *Dictionary of psychology* (revised ed.). New York, NY: Dell.
- Clarizio, H., & Bernard, R. (1981). Recategorized WISC-R scores of LD children and differential diagnosis. *Psychology in the Schools*, 18(1), 5-12.
- Dash, U., Dennis, S., Mueller, H., Mancini, G., Snart, F., & Short, R. (1983). WISC-R subtest variability in a sample of clinic-referred Canadian children. *Canadian Journal of Behavioral Science*, 15(3), 211-227.
- Dean, R. (1977). Patterns of emotional disturbance on the WISC-R. *Journal of Clinical Psychology*, 33(3), 486-490.
- Dean, R. (1978). Distinguishing learning disabled and emotionally disturbed children on the WISC-R. *Journal of Consulting & Clinical Psychology*, 46(2), 381-382.
- Dinning, W. (1976). Variance components of the WISC-R in eleven age groups. *Psychological Reports*, 38, 1001-1002.
- French, J. (1966). The logic of and assumptions underlying differential testing. In A. Anastasi (Ed.), *Testing problems in perspective*. Washington, DC: American Council on Education.
- Groff, M., & Hubble, L. (1981). Recategorized WISC-R scores of juvenile delinquents. *Journal of Learning Disabilities*, 14(7), 519-520.
- Guilford, J. (1954). *Psychometric methods* (2nd ed.). New York, NY: McGraw-Hill.
- Gutkin, T. (1978). Some useful statistics for the interpretation of the WISC-R. *Journal of Consulting & Clinical Psychology*, 46(6), 1561-1563.

- Gutkin, T. (1979a). WISC-R scatter indices: Useful information for differential diagnoses? *Journal of School Psychology*, 17(4), 368-371.
- Gutkin, T. (1979b). The WISC-R Verbal Comprehension, Perceptual Organization, and Freedom from Distractibility deviation quotients: Data for practitioners. *Psychology in the Schools*, 16(3), 359-360.
- Hale, R. (1979). The utility of WISC-R subtest scores in discriminating among adequate and underachieving children. *Multivariate Behavioral Research*, 14(2), 245-253.
- Hale, R., & Landino, S. (1981). Utility of WISC-R subtest analysis in discriminating among groups of conduct problem, withdrawn, mixed, and nonproblem boys. *Journal of Consulting & Clinical Psychology*, 49(1), 91-95.
- Haynes, J., & Bensch, M. (1981). The P>V sign on the WISC-R and recidivism in delinquents. *Journal of Consulting & Clinical Psychology*, 49(3), 480-481.
- Hirshoren, A., & Kavale, K. (1976). Profile analysis of the WISC-R: A continuing malpractice. *Exceptional Children*, 23(1), 83-87.
- Karnes, F., & Brown, K. (1980). Factor analysis of the WISC-R for the gifted. *Journal of Educational Psychology*, 72(2), 197-199.
- Kaufman, A. (1975). Factor analysis of the WISC-R. *Journal of Consulting & Clinical Psychology*, 43(1), 135-147.
- Kaufman, A. (1979). *Intelligent testing with the WISC-R*. New York, NY: J. Wiley & Sons.
- Kaufman, A. (1980). Issues in psychological assessment: Interpreting the WISC-R intelligently. In B. Lahey & A. Kazdin (Eds.), *Advances in clinical child psychology*. New York, NY: Plenum.
- Kaufman, A. (1981). The WISC-R and learning disabilities assessment: State of the art. *Journal of Learning Disabilities*, 14(9), 520-526.
- Kelley, T. (1923). A new method for determining the significance of differences in intelligence and achievement scores. *Journal of Educational Psychology*, 14(6), 321-333.
- Kunce, J., & McMahon, R. (1979). Neuropsychological significance of children's Wechsler intelligence scores. *Psychological Reports*, 44(3, pt. 1), 787-790.
- Loro, B., & Woodward, J. (1976). Verbal and performance IQ for discrimination among psychiatric diagnostic groups. *Journal of Clinical Psychology*, 32(1), 107-114.
- Lutey, C. (1977). *Individual intelligence testing: A sourcebook*. Greeley, CO: Lutey.
- Matheson, D., Mueller, H., & Short, R. (in press). The validity of Bannatyne's Acquired Knowledge category as a separate construct. *Journal of Psychoeducational Assessment*.
- McIntosh, W. (1974). The use of a Wechsler subtest ratio as an index of brain damage in children. *Journal of Learning Disabilities*, 7(2), 161-163.
- Miller, M. (1980). On the attempt to find WISC-R profiles for learning and reading disabilities. *Journal of Learning Disabilities*, 13(6), 52-54.
- Mollenkopf, W. (1950). Predicted differences and differences between predictions. *Psychometrika*, 15(4), 409-417.
- Mosier, C. (1951). Batteries and profiles. In E. F. Lindquist (Ed.), *Educational measurement*. Washington, DC: American Council on Education.
- Mueller, H., Matheson, D., & Short, R. (1983). Bannatyne-recategorized WISC-R patterns of mentally retarded, learning disabled, normal, and intellectually superior children: A meta-analysis. *The Mental Retardation and Learning Disability Bulletin*, 11(2), 60-78.
- Nunnally, J. (1967). *Psychometric theory*. New York, NY: McGraw-Hill.
- Piotrowski, R., & Grubb, R. (1976). Significant subtest score differences on the WISC-R. *Journal of School Psychology*, 14(2), 202-206.
- Quattrocchi, M., & Sherrets, S. (1980). WISC-R: The first five years. *Psychology in the Schools*, 17(3), 297-312.
- Reschley, D. (1978). WISC-R factor structures among Anglos, Blacks, Chicanos, and native American Papagos. *Journal of Consulting & Clinical Psychology*, 46, 417.

- Reynolds, C. (1981). A note on determining significant discrepancies among category scores on Bannatyne's regrouping of WISC-R subtests. *Journal of Learning Disabilities*, 14(8), 468-469.
- Reynolds, C., & Gutkin, T. (1980). Stability of the WISC-R factor structure across sex at two age levels. *Journal of Clinical Psychology*, 36(3), 775.
- Ryckman, D. (1981). Searching for a WISC-R profile for learning disabled children: An inappropriate task? *Journal of Learning Disabilities*, 14(9), 507-526.
- Salvia, J., & Ysseldyke, J. (1981). *Assessment in special education and remedial education* (2nd Ed.). Dallas, TX: Houghton-Mifflin.
- Sattler, J. (1982). *Assessment of children's intelligence and special abilities* (2nd Ed.). Boston, MA: Allyn & Bacon.
- Segel, D. (1934). *Differential diagnosis*. Baltimore, MD: Warwick & York.
- Smith, G. (1970). *A simplified guide to statistics for psychology and education* (4th ed.). New York, NY: Holt, Rinehart & Winston.
- Tew, B. (1977). Spina Bifida children's scores on the WISC. *Perceptual & Motor Skills*, 44(2), 381-382.
- Thompson, A. (1981). Some counter thinking about learning disabilities. *Journal of Learning Disabilities*, 14(7), 394-396.
- Thompson, R. (1980). Diagnostic utility of WISC-R measures with children referred to a developmental evaluation center. *Journal of Consulting & Clinical Psychology*, 48(4), 440-447.
- Tierney, R., & Ames, W. (1978). An examination of the diagnostic claims of the Illinois Test of Psycholinguistic Abilities. *Journal of Learning Disabilities*, 11(9), 586-590.
- Vance, H., & Singer, M. (1979). Recategorization of the WISC-R subtest scaled scores for learning disabled children. *Journal of Learning Disabilities*, 12(8), 487-490.
- Vance, H., Wallbrown, F., & Blaha, J. (1978). Determining WISC-R profiles for reading disabled children. *Journal of Learning Disabilities*, 11(1), 55-58.
- Van Hagen, J., & Kaufman, A. (1975). Factor analysis of the WISC-R for a group of mentally retarded children and adolescents. *Journal of Consulting & Clinical Psychology*, 43, 661-667.
- Wallbrown, F., Vance, H., & Blaha, J. (1979). Developing remedial hypotheses from ability (WISC-R) profiles. *Journal of Learning Disabilities*, 12(1), 59-63.
- Webster, R., & Lafayette, A. (1980). Distinguishing among three subgroups of handicapped students using Bannatyne's recategorization of the WISC-R. *Journal of Educational Research*, 73(4), 237-240.
- Wechsler, D. (1974). *WISC-R manual*. New York, NY: Psychological Corporation.

PETER EVANS

NORMA MICKELSON

and

NELSON C. SMITH

University of Victoria

Computer-Assisted Instruction in College Composition

This research investigated the use of microcomputers in the development of prose writing ability of university students. A computer-assisted language program employing the Apple II computer in a tutorial function was used in the study. The program presented information, responded to answers, and evaluated student input with a programmed response.

Two groups of 25 freshman students enrolled in basic composition courses at the University of Victoria were used as subjects. Student performance was measured through pre- and post-tests of sentence structure, error analyses (tabulated by the computer), and an examination of sentence types used in written compositions.

The results indicated few significant differences in performance between control and experimental groups. All differences, however, favored the experimental group. It was concluded that computer-assisted instruction appears to be at least as effective as regular instruction. This suggests that computer-assisted instruction can be used to supplement face-to-face teaching, thus freeing the instructor for individual attention to student needs.

Introduction

The objective of this research was to investigate the use of microcomputers in the development of prose writing ability at the university level. The research was

Peter Evans is a Professor of Language Arts at the University of Victoria. He has published numerous articles on aspects of language learning as well as collaborated in the writing of student texts for the junior high school reading program.

Norma Mickelson is a Professor of Education at the University of Victoria with extensive experience in public schools, both as a teacher and as a supervisor of instruction. The current focus of her work is in the application of learning principles to classroom practice.

Nelson C. Smith, Associate Professor of English at the University of Victoria, has been serving as Director of First Year English and the Language Programme for the past two years.

heuristic in nature, its purpose being to examine the potential of computer technology in the development of student writing achievement and to ascertain directions for future research and application.

The computer-assisted language program used in this study was developed to function as a "tutor." Apple II discs were used. The program presented information, responded to answers, and evaluated student input with a programmed response. While it was a right/wrong format and thus did not evaluate the quality of the errors, it did make the student aware of mistakes and also supplied correct responses. Error detection by the computer enabled the student to assume responsibility for analyzing and correcting mistakes. The pupils, however, if they so desired, could refer to the instructor for discussion and explanation of errors. This approach seemed particularly appropriate at the college level where student-directed learning is one goal of instruction and where remedial activities appear to be necessary for improvement of writing skills.

Since student attitude to writing is a variable which may affect composition, the researchers used the Writing Apprehension Measure (WAM) (Daly & Miller, 1975b) to assess student attitudes. Daly hypothesized that high writing apprehension scores correlate negatively with success in writing and also predicted that females would score lower in writing apprehension than would males. The Writing Apprehension Measure is a 26-item Likert scale with a range of five responses. Items represent anxiety about writing. The scale identifies highly apprehensive students who find writing unrewarding and anxiety-producing at one end of the continuum, and at the other, less apprehensive individuals who are positive and confident in their attitudes towards writing.

In summary, then, the study addressed three basic questions:

1. Does computer-assisted instruction affect writing performance?
2. What are student attitudes to computer-based instruction?
3. Does attitude towards writing affect performance?

Review of the Literature

Readings in the application and implementation of the microcomputer in the language arts consist mainly of position papers and theoretical constructs. The scarcity of empirical studies reflects the short history of computers in the classroom and the more frequent application of computer-assisted instruction to the mathematics and science areas.

Arguments for and against computer-assisted language learning were articulated by Spolsky (1970) in a discussion of the nature of language learning and the theoretical basis of grammar teaching using a computer. Spolsky argued that the computer was a valuable tool because it forced the student-operator to develop a rationale for his problem-solving, thereby enhancing his achievement.

Gagné (in Stotsky, 1982) saw the computer as having potential for the teaching of basic skills in language arts classes. He discussed the need for bringing basic skills to the point of automaticity in order to supplement the limited capacity of short-term memory in problem-solving. He pointed out the utility of the computer in providing consistency in the drill and practice techniques needed to bring subordinate skills to a level of automaticity.

Many authorities in language education have stressed the need for instruction in grammar, spelling, word definition, and sentence construction to ensure a reason-

able level of composition ability and literacy. Others have pointed to the potential of the computer to motivate students and make routine material interesting. The importance of increased motivation was specifically mentioned by Holmes and Kidd (1982). The tutorial method of computer-assisted instruction seemed to them to be an effective use of the microcomputer's unique characteristics. Interaction between the learner and the material is possible when "branching" is used to allow for individual pacing and reinforcement (Holmes & Kidd, 1982).

Boyd, Keller, and Kenner (1982) felt that grammatical constructions lend themselves readily to effective computer-assisted instruction by means of multiple choice items.

The teaching of grammar and its effects on writing ability continue to be discussed in the literature. Stotsky (1975) pointed to the failure of past studies of grammar usage to test for skills taught. In a review of the research she included the work of Mellon (1982) and O'Hare (1973) which showed an improvement in T-unit measures (Loban, 1963) and an increase in variety of sentence structures as a result of instruction.

Although a consensus of opinion in the literature suggests that the teaching of grammar in a traditional sense does not necessarily result in improvement of writing skills, it seems possible that using a computer as a novel approach to teaching basic grammar skills might reverse this situation.

Design

The Research Model

A review of the literature reveals two fundamental factors which impinge upon the present research. Firstly, there are mixed opinions and conflicting evidence regarding the impact upon composition of the teaching of fundamental grammatical concepts such as sentence types and coordination and subordination of sentence elements. Secondly, there is little evidence of the efficacy of computer-based instruction upon the learning of grammatical concepts. Added to these two factors are concerns about other antecedent conditions surrounding the teaching and learning of composition skills. These include student attitudes to writing and the differential influence of sex upon performance.

In order to examine the interrelationships of these factors, this study was organized to examine the instrumental effects of learning fundamental grammar concepts upon composition. Further, two approaches to the development of these concepts were compared. Finally, the effects of attitude toward writing and the sex influence on performance were investigated. The research design employed was the Static Group Comparison Model (Campbell & Stanley, 1963, p. 8). This model permits comparison of the performance of two intact groups, one of which receives the experimental treatment.

Subjects and Procedures

Two groups of 25 freshman students enrolled in basic composition courses at the University of Victoria were used as subjects. All freshman students had been randomly assigned to 16 sections of the course. The two groups used in the research were selected on the basis of scheduling availability. Through attrition (dropping of the course) four were lost from one group and three from the other.

The control group (RI) was given regular classroom instruction in sentence structure. The experimental group (CAI) was given similar instruction using an

Apple II computer for three one-hour sessions. Both instructional programs presented, described, and classified sentence types. Clauses were identified and bracketed; sentence errors and effective grammatical devices were identified by example and definition. For the control group (RI), these definitions were presented by the teacher in traditional classroom instruction. Traditional classroom instruction in this study involved the teaching of basic writing skills including the knowledge of grammatical structures and rhetorical devices through face-to-face presentation. Concepts were presented and defined. This activity was followed by student application in assigned writing activities. For the experimental group, an Apple II Disc (tutorial) led the student through a series of activities designed to cover the same material as in-class instruction. Both programs included simple sentences, complex sentences, compound sentences, effective use of clauses, subordination, coordination, parallelism, and sentence errors. The student's task was to select ways to correct faulty sentences.

Instrumentation

Student performance was measured as indicated in Table 1.

Variables 1 and 2, the Objective Test, measured the students' knowledge of the specific concepts taught only in the computer tutorial program. This 40-item exercise was a pencil and paper test, and the pretest was used as the covariate in the statistical analysis.

Variables 3 and 4, the Essay Marks, were a general assessment of the quality of the students' writing. The relationship between the learning of grammatical concepts and the quality of writing was examined. Variables 5 to 14 and 23 to 28, sentence types and rhetorical devices, were used as a specific analysis of the degree to which the treatment and control learning activities influenced the use of these structures in writing. Variables 5-14 were pre-treatment results and variables 23-28 were post-treatment results.

Variables 17 and 18, Error Analysis, were the results of the computer tabulation of student performance on the computer tutorial activities. These variables indicated the degree of difficulty students had with the lesson material.

Variables 20, 21 and 22, Course Marks and Marks on Sentence Structure in the Final Exam, were performance values assigned by instructors for work done in the course. These evaluations were examined to determine the relationships between specific computer learning activities and general performance in the course.

Finally, student attitudes were measured by Daly's Writing Apprehension Measure and the results correlated with the grammar and composition scores.

Analysis

An analysis of covariance was used to analyze the data with the pre-treatment grammar scores as the covariate. The level of significance selected for the study was .05.

Results

A. Experimental Results

Results indicate few significance differences in performance between control and experimental groups. Both groups, however, improved over time. Computer-assisted instruction, therefore, appears to be no less effective than regular instruction.

TABLE 1
ANALYSIS OF COVARIANCE

Variable Number	Variable Name	Male/ Female	CAI/ RI	WAM Score Low/High
VO1	Objective Pre-Test	NS	NS	NS
VO2	Objective Post-Test	NS	NS	S
VO3	Essay Mark: Pre-Experimental	NS	NS	NS
VO4	Essay Mark: Post-Experimental	S	NS	NS
VO5	Essay—No. of Simple Sentences: Pre-Exp.	NS	NS	NS
VO6	Essay—No. of Compound Sentences: Pre-Exp.	NS	NS	NS
VO7	Essay—No. of Complex Sentences: Pre-Exp.	NS	S	NS
VO8	Essay—Total No. of Sentences: Pre-Exp.	NS	NS	NS
V13	Essay—No. of Sentence Errors: Pre-Exp.	NS	S	S
V14	Essay—No. of Effective Devices: Pre-Exp.	NS	NS	NS
V17	Error Analysis I CAI	NS		NS
V18	Error Analysis II CAI	NS		S
V20	Final Course Mark	NS	NS	NS
V21	Mark on Exam Sentence Structure	NS	NS	NS
V22	Final Exam Mark	NS	NS	NS
V23	Essay—No. of Simple Sentences: Post-Exp.	NS	S	NS
V24	Essay—No. of Compound Sentences: Post-Exp.	NS	NS	NS
V25	Essay—No. of Complex Sentences: Post-Exp.	NS	NS	NS
V26	Essay—Total No. of Sentences: Post-Exp.	NS	NS	NS
V27	Essay—No. of Sentence Errors: Post-Exp.	NS	NS	NS
V28	Essay—No. of Effective Devices: Post-Exp.	NS	NS	NS

NS = No Significance
S = Significant Difference: See Table 2

An examination of the data reveals the following:

1. On the post-treatment objective test, the low-apprehensive group scored higher than the high-apprehensive group (Table 2: V02).
2. The mean percentage for females in the post-treatment essay was higher than for males.
3. The CAI group used more complex sentences than the RI group in the pre-treatment compositions. This difference between groups was not present in the essays written after treatment.
4. The CAI group made more sentence errors (comma splice, ineffective subordination, ineffective coordination and parallelism) than the RI group in the pre-treatment essays; post-test differences were not significant. The group that scored low on the Writing Apprehension Measure made more sentence errors in the pre-treatment essay test than did the high-scoring WAM group. On the post-treatment essay test the low-scoring group made fewer errors than did the higher-scoring WAM group.

- 5. CAI subjects having a low WAM score made more errors in the computer program than did those having a high WAM score.
- 6. The CAI group used fewer simple sentences in the post-treatment essay than did the RI group.

TABLE 2
SIGNIFICANCE DIFFERENCES

Variable	Male	Female	CAI	RI	Signif. of F	WAM Score		Signif. of F
						Low	High	
VO2 Objective Post-Test	—	—	—	—	—	25.78	21.73	.033
VO4 Essay Mark: Post-Experimental	64.5	69.35	—	—	.025	—	—	—
VO7 Essay—No. of Sentence Errors: Pre-Experimental	—	—	9.10	5.25	0.11	—	—	—
V13 Essay—No. of Sentence Errors: Pre-Experimental	—	—	5.48	2.75	.020	6.18	2.92	.048
V18 Error Analysis II CAI	—	—	—	—	—	62.7	50.0	.041
V23 Essay—No. of Simple Sentences: Post-Experimental	—	—	5.64	8.65	.037	—	—	—

TABLE 3
OBJECTIVE TEST MEAN SCORES

	Pre-Test	Post-Test
CAI	16.9	21.8
RI	16.4	22.2

B. Anecdotal Results

In analyzing student reactions (as recorded on individual log sheets) to the computer-assisted instruction, the researchers noted that students liked the individual attention provided, the chance to set their own pace, and the immediate feedback given in the program. Mentioned also was the lack of distraction. Many comments suggested the need for frequent loops to allow for referral back in the program for information, the usefulness of frequent repetition to provide reinforcement, and the desirability of being able to start or stop midway through the program. One student with previous experience with computers recommended more incentives or rewards to eliminate boredom. Students also expressed interest in knowing their scores because they thought this would help them evaluate their progress. Two subjects mentioned frustration when they did not understand why their answers were incorrect. In general, females appeared to be more apprehensive about using computers than did males.

Discussion

While at first glance there appear to be few differences in effect between regular and computer-assisted instruction, a detailed analysis of the data reveals noteworthy differences favoring the treatment (CAI) group. These differences are discussed below under each of the three basic questions considered in the study.

A. Does computer-assisted instruction affect writing performance?

1. Results indicate a significant difference between the CAI group and the RI group in use of complex sentences in the pre-treatment essay. This difference was not noted in the post-test essay. While this may appear to be a reduction in the quality of performance, an examination of the essays revealed that this result indicated, in fact, a reduction of run-on sentences in the compositions of the CAI group. It should be noted that in the original design of the study, the analysis did not include run-on sentences. These were categorized as complex sentences. The results appear to indicate increased awareness by the students of the importance of avoiding run-on sentences.
2. The result showing a significance difference in essay writing following instruction in favor of females in both groups suggests that in both computer-assisted instruction and regular instruction, females improved significantly. This result would tend to contradict the pervasive assumption that women do not benefit from technological approaches to learning.
3. Significant differences in the number of simple sentences used in the post-test favoring the regular instruction group suggest that students showed less improvement in the use of a variety of sentence types than did those in the computer-assisted group. Thus the tutorial program may have left a more powerful impression on the learners as to the variety of sentence forms and structures available for use in written composition than did regular classroom instruction.

B. What are students' attitudes to computer-based program instruction?

1. In general, most students reacted positively to computer-assisted instruction as evidenced by comments such as, "It's fun" and "I can set my own pace."

2. Students noted the need for an explanation of incorrect answers in the computer tutorial. This observation reinforces the importance of teacher participation in learning and suggests that computers might best be used as an adjunct to regular classroom instruction.
3. Many of the disadvantages of computer-assisted instruction expressed by students could be eliminated in more sophisticated programs with frequent loops for optional reviewing, reinforcement, or further explanation. A summary of the student's correct and incorrect answers could also be provided with areas of difficulty pinpointed.

C. Does attitude towards writing affect performance?

In the post-treatment objective grammar test, students with low WAM scores performed better than did those who were more apprehensive, a finding consistent with Daly's hypothesis. In essay writing, high apprehensive students appeared to be more concerned about mechanics than did low apprehensive students. Low apprehensive students appeared to profit more from instruction as indicated by sentence error data. The result showing that low apprehensive students made more errors in the computer tutorial suggests that they appear to be willing to risk making errors to a greater degree than do students who are highly apprehensive. More research is needed in this area.

Conclusions and Recommendations

The results of this study suggest that computer-assisted instruction appears to be no less effective than regular classroom teaching of sentence mechanics. The fact that there were no measurable differences in performance of students on most variables indicates that computer-assisted instruction may produce results as good as face-to-face instruction. In no case did subjects who were exposed to the microcomputer in place of regular instruction achieve less well than did those who attended regular classes. In fact, there appeared to be several important benefits for those who used the computer. For example, students developed both greater ability to use variety in sentence structure and accuracy in construction of sentences. Outcomes such as these are considered to be fundamental to a program of English composition instruction.

Computer-assisted programs of instruction in fundamental composition skills appear to be both effective and viable. Such programs take less teacher time than regular classroom instruction and thus might effectively supplement and vary approaches taken in the teaching of English. The more factual and technical aspects of course content with the focus on student responsibilities for learning, for example, could be presented by computers, while the application of such content might be handled by human instructors. Remedial teaching could well be handled, at least in part, using computer-assisted instruction.

Specific recommendations arising out of the study are as follows:

1. There were certain problems in logistics during the experimental program. These involved the availability of microcomputers and the scheduling of appropriate practice sessions. The study should be replicated in such a manner as to obviate these problems.
2. The identification of errors and sentence structures in essays should be more rigidly controlled.

3. Students in first-year English should be given an opportunity to supplement the learning of fundamental composition skills through computer-assisted instruction.
4. Females, in particular, should be encouraged to participate in computer-aided courses of instruction.
5. The writing apprehension level of students might be investigated further as a criterion in planning learning activities.
6. Studies might be conducted to determine how increased exposure to computer programs affects learning rate.

References

- Adams, E. N., Morrison, H. W., & Reddy, J. M. (1968). Conversations with a computer as a technique of language instruction. *The Modern Language Journal*, 42, 3-15.
- Boyd, G., Keller, A., & Kenner, R. (1982). Remedial and second language English teaching using computer-assisted learning. *Computers and Education*, 6, 105-112.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.
- Daly, J. A. (1977). The effects of writing apprehension on message encoding. *Journalism Quarterly*, 54, 566-572.
- Daly, J. A., & Miller, M. D. (1975a). Apprehensions of writing as a predictor of message intensity. *Journal of Psychology*, 89, 556-572.
- Daly, J. A., & Miller, M. D. (1975b). The development of a measure of writing apprehension. *Research in the Teaching of English*, 9, 242-249.
- Daly, J. A., & Shamo, J. (1975). Writing apprehension and occupational choice. *Journal of Occupational Psychology*, 49, 55-56.
- Holmes, G., & Kidd, M. E. (1982). Second-language learning and computers. *The Canadian Modern Language Review*, 3, 503-516.
- Loban, W. (1963). *The language of elementary school children*. Urbana, IL: (NCTE Research Report No. 1).
- Mellon, J. C. (1982). *Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition*. Cambridge: Harvard University.
- O'Hare, F. (1973). *Sentence-combining: Improving student writing without formal grammar instruction*. Urbana, IL: (NCTE Research Report No. 15).
- Spolsky, B. (1970). The language teacher in the computer age. In R. Lugton (Ed.), *Preparing the E.F.L. teacher*. Philadelphia: The Center for Curriculum Development.
- Stotsky, S. (1975). Sentence-combining as a curricular activity: Its effect on written language development and reading comprehension. *Research in the Teaching of English*, 9, 30-71.
- Stotsky, S. (1982). Developments in learning psychology: Implications for instructional design: Effects of computer technology on instructional design and development. An interview with Robert M. Gagné, in *Educational Technology*.

WAYNE K. HOY

Rutgers University

and

DAVID A. SOUSA

West Orange Public Schools

Delegation: The Neglected Aspect of Participation in Decision Making¹

Data were collected from the principals and teachers in 55 secondary schools. Using the Aston methodology, an index of shared decision making was developed, and three hypotheses were derived and tested. All were supported. Principals who shared decision making with subordinates had more loyal teachers, had more satisfied teachers, and had teachers who described the authority structure as less rigid than principals who did not share decision making. The research underscored an important theoretical distinction between sharing decision making authority in terms of delegating and sharing in terms of joint decision making.

The issue of participation in decision making remains a popular topic in organizational analysis; in fact, subordinate participation is a basic tenet of human relations theory (Likert, 1967; Strauss & Rosenstein, 1970). Despite its common usage, there is relatively little consensus on the nature and meaning of participation. Schregle (1970, p. 117) cogently comments, "Workers' participation has become a magic word in many countries. Yet almost everyone who employs the term thinks of something different." Although Shregle's observation seems a bit strong, participation in decision making clearly suffers from a variety of uses (Dachler & Wilpert, 1978). For example, Likert (1967) views participation as a management style; Leavitt (1965) and Tannenbaum (1974) see it as equalization of power or sharing power; and Davis (1967) relates it to group involvement and group decision making. Moreover, some theorists distinguish between participation in decision making and

Wayne K. Hoy is Professor and Associate Dean for Academic Affairs, Graduate School of Education, Rutgers University. He is coauthor of *Educational Administration: Theory, Research and Practice*.

David A. Sousa is Director of Science, West Orange Public Schools, West Orange, New Jersey.

delegation of decision making (Davis, 1963; Locke & Schweiger 1979); others do not (Lowin, 1968; Sashkin, 1976).

In the present investigation the concern was with examining one aspect of participation in decision making—the extent to which principals delegate decisions to subordinates. The approach was first to develop an index of the principal's propensity to share important decisions with subordinates and then to examine the relationships of that propensity with teachers' loyalty to principal, teachers' satisfaction with their jobs and teachers' perception of the degree of hierarchical control.

Conceptual Perspective

Four specific concepts provided the perspective employed in the study; each is identified and described below.

Participation in Decision Making is a broad term often used to refer either to joint decision making or to the process of delegating specific decisions to subordinates. The process of delegating decision authority to subordinates, however, is not the same as jointly arriving at a decision. In joint decision making the participants are involved in the *process* of making decisions; they share a common effort. *Delegation of decision making* is the entrusting of authority to others; an administrator assigns specific decisions to others, usually subordinates. The subordinates are then free to make their decisions without further consultation with the supervisor. Many studies have centered on participation in decision making but few in recent years have examined the benefits of delegated decision making.

Principals who delegate recognize that they are yielding some of their authority and that, for the most part, they are bound by the subordinates' decisions. What are the benefits? Does such a practice improve job satisfaction among subordinates? Does it increase subordinate loyalty? Does it alter teachers' perceptions of the authority structure of the school? These were the basic questions that guided the empirical phase of this study.

Hierarchy of Authority is the extent to which the focus of decision making is prestructured by the formal authority system. Organizations vary in the degree to which decisions must be first approved by superiors before any action can be taken. School organizations in which teachers are given little or no freedom to accomplish assigned tasks unimpeded by their superiors have a rigid hierarchy of authority. What we are calling hierarchy of authority has also been termed tightness or closeness of supervision by other researchers (Blau & Scott, 1962, pp. 145-150).

Job Satisfaction reflects a sense of gratification with one's position of employment in the organization; it is the extent to which an individual is satisfied with such things as authority, relative position compared to other workers, opportunities for professional growth, recognition and acceptance granted by superiors, and the degree to which the job is consistent with one's career expectations. Aiken and Hage (1966) call the lack of satisfaction with one's job alienation from work; this study uses the positive term, job satisfaction, because it more accurately reflects the meaning of the operational definition and it avoids the questionable comparison of the concept with Marx's alienation from the process of production (Schacht, 1970).

Subordinate Loyalty to Superior was first introduced by Blau and Scott (1962) as an important variable mediating work relations in hierarchical organizations, but despite its apparent theoretical significance, they simply noted that superiors who command loyalty are "liked," "accepted," and "respected." Murray and Corenblum (1966) defined the term more rigorously, arguing that, in addition to the affect-

based definition of Blau and Scott (1962), at least two other dimensions of subordinate loyalty—cognitive and behavioral—should be considered. Subordinates who are cognitively loyal manifest an unquestioning faith and trust in the superior, and those who are behaviorally loyal show an actual or expressed willingness to follow when the superior accepts a new position. Subsequent research in schools (Hoy & Rees, 1974; Hoy & Williams, 1971) demonstrated that the three dimensions of subordinate loyalty vary concomitantly; hence, *loyalty to the principal* is the extent to which the principal is liked, respected, accepted, trusted and followed by teachers.

The research, then, examined the relationships between delegated decision making and the three concepts of (1) loyalty, (2) job satisfaction, and (3) hierarchy of authority.

Some Hypotheses

Principals are required to make many decisions in the course of their work. Some decisions are trivial while others have a major impact on the operation of the school. Principals have devised various methods for dealing with decision making. Those who make decisions without input from other parts of the organization have been referred to as authoritarian. At the other extreme are those who seek extensive input before deciding. These are often referred to as democratic or participative administrators. In his study, Cross (1980) noted that principals who sought input relied on one or more of the following four sources of information when making decisions: subordinates, extraordinates (people not directly affiliated with the school), superordinates, and peers. He found that although principals rarely consulted superordinates, they were strongly influenced in their decisions by subordinates.

Teachers in recent years have requested greater participation in decisions that affect them and in which they have some expertise. The delegation of such decisions by principals to subordinates is likely to influence their perceptions of the control structure of the school. In a study of social welfare agencies, Aiken and Hage (1966) found a substantial negative correlation between the degree of subordinate participation in decision making and extent of hierarchical control. Moreover, Hoy and his colleagues (1983) report the same results in their comprehensive study of principals and teachers in 41 high schools in New Jersey; the greater the degree of teacher participation in decision making, the less the sense of hierarchy of authority. In both studies, however, participation in decision making was defined in broad terms. In each case subordinates were asked such questions as, "How frequently do you participate in the decision to hire new staff?" and "How frequently do you participate in decisions on the adoption of new policy?" It is unclear, however, whether participation meant the act of being consulted by superiors, or jointly making the decisions with superiors, or making decisions on their own that had been delegated to them by superiors. Delegating the decision to subordinates has the potential to maximize participation; the principal entrusts the decisions to subordinates. The extent to which such delegation occurs seems likely to affect the teachers' perception of the formal authority structure of the school; consequently it was hypothesized that:

H.1. The greater the propensity of the principal to delegate decision making, the less hierarchy of authority perceived by teachers.

Teachers do not want to be involved in all kinds of decisions. Kunz and Hoy (1976) found that school decisions were characterized by a "zone of acceptance,"

and that teachers willingly left decisions within this zone to administrators. In an extensive three-year study, Crockenberg and Clark (1979) were able to get active teacher participation only after formal mechanisms were established through faculty constitutions and faculty councils. Moreover, they found that the decision-making areas of particular concern to teachers were distinctly instructional rather than administrative. Duke, Showers, and Imber (1980) found that most teachers were willing to participate in decision making only when such involvement meant actual influence over decisions, and that these decisions had great potential for improving classroom life and student outcomes. Healey (1980), in a survey of 150 elementary principals in Illinois, also found that the more the decisional task related to classroom instruction, the more faculty involvement in decision making occurred.

Job satisfaction reflects a sense of satisfaction with one's career, professional development, authority, position, and ability to fulfill professional responsibilities. Aiken and Hage (1966) referred to such lack of satisfaction with one's job as work alienation, and several studies (Aiken & Hage, 1966; Hoy et al., 1983) found that the lack of participation by subordinates in organizational decision making was strongly related to their alienation from work. Conway's (1976) research showed that secondary school teachers were more satisfied with their jobs if they participated in and influenced decisions affecting instruction, and if such participation was not excessive. The Isherwood and Taylor (1978) study of 48 secondary schools in Quebec found that teacher participation in decision making through school councils had a positive effect on their job satisfaction. In their extensive review of studies on decision making in non-educational organizations, Locke and Schweiger (1979) concluded that a majority of their studies supported the premise that increased participation in decision making resulted in increased job satisfaction. In schools, if teachers are granted authority to make important decisions, then they should feel a greater sense of self-attainment and influence, thereby increasing their satisfaction with their jobs; hence, it was hypothesized that:

H.2. The greater the propensity of the principal to delegate decision making, the greater the job satisfaction of teachers.

Many principals have only the power and authority of their office. In a real sense they are merely officers or bureaucrats, not leaders. How can principals increase the base of their authority? The informal organization of teachers is an important source of authority, and principals need to find ways to tap the informal authority network. Where formal authority is legitimated by the values that have become routinized in legal contracts, informal authority is legitimated by the common values and sentiments that emerge in the work group, *particularly by the loyalty the superior commands from group members* (Blau & Scott, 1962, p. 144). The importance of teacher loyalty to principals seems clear. Those principals who can command loyalty from teachers have an advantage in enlarging their authority base by establishing effective informal authority over teachers (Hoy & Rees, 1974, p. 269).

Several studies have shown that rigid control was negatively related to subordinate loyalty to the supervisor. Blau and Scott (1962) provided original evidence that tight control in social welfare agencies was negatively related to commanding subordinate loyalty. Similarly, two studies in secondary schools demonstrated that the more authoritarian the behavior of the principal, the less the teacher loyalty to the principal (Hoy & Rees, 1974; Hoy, Tarter, & Forsyth, 1978). Moreover, in a study of secondary school principals, a global measure of participa-

tion in decision making was found to be positively related to teacher loyalty to the principal (Hoy, Newland, & Blazovsky, 1977). When a principal voluntarily delegates important decisions to a teacher or group of teachers, it reflects the principal's confidence in the ability of teachers to reach acceptable, reasonable, and useful decisions. Such displays of confidence and trust should increase the sense of loyalty teachers feel toward the principal; therefore, it was hypothesized that:

H.3. The greater the propensity of the principal to share decision making, the greater the degree of teacher loyalty to the principal.

Design and Method

In order to test the hypotheses of this study, an index of propensity to delegate decision making was developed for a sample of secondary school principals; and the perceptions of the teachers in each school were analyzed.

Sample

Data were collected on 55 public senior high schools located in 11 counties in New Jersey. In selecting the schools, an attempt was made to obtain a variety of community types including urban, suburban, and rural areas. The size of the high schools in the sample ranged from 339 to 2400 students (with a mean of 1251 students), and the full-time professional staff ranged from 34.5 to 168 (with a mean of 97.0 professionals per school). Although no claim is made that the sample of schools was random, it seems to be a representative sample of New Jersey secondary schools, especially with respect to size. If any type of school was underrepresented, it was probably the large urban-core high school—only two such schools were included in the sample.

Independent Variable—Propensity to Delegate

In conducting this study, it was necessary to select a set of decisions that were typical for public school settings, that were relevant and important, and that discriminated among schools in terms of the principal's decision-making authority. The Aston approach developed by Pugh and Hickson (1978) and their colleagues was deemed appropriate for the task.² This approach analyzes the structure of organizations through interviews with the chief executives. The interview schedule, designed to gather more objective data about the organization, was administered to each principal. One portion of the schedule contained a series of school-related decisions that described the principal's autonomy. In effect, the autonomy items measured the number of decisions that were made by building principals without the immediate approval of their superordinates. Those decisions that were autonomous to principals represented the maximum number of decisions that principals in each building could legitimately delegate to subordinates. Specifically, principals were asked to identify the level within the school system at which 22 different decisions were made. The six decision making levels from which they could select were: teacher, department head, assistant principal, principal, superintendent, and local board of education. It was explained that decision making authority meant action could be taken on the decision, even though the decision might be subject to routine ratification at a later time. This meant that, although others might ratify the decision, its intentions would not be altered.

The principals' responses were subjected to item analysis, and the ten decision questions listed below were found to discriminate among the schools.

Who decides:

1. The number of department heads in the schools?
2. The appointment of a teacher?
3. The amount of money to be allocated to each department?
4. How unallocated funds will be spent?
5. Whether a new course or subject will be introduced?
6. Whether a new program will be introduced?
7. The procedures for purchasing materials for the school?
8. Which suppliers or materials will be used?
9. To create a new department?
10. To create a new teaching or administrative position?

By identifying the level at which decisions were made for each school, the principals provided a means for investigating the extent to which they delegated their decision-making authority. Specifically, the comparison was made by establishing a ratio in which the numerator represented the number of decisions made by the subordinates (teachers, department heads, or assistant principals) in the school, and the denominator represented the total number of within-school decisions made. This ratio was called the delegation of decision making index and was expressed by the equation:

$$\frac{\text{Number of subordinate decisions}}{\text{Total number of within-school decisions}} = \text{Delegation of decision making index}$$

The index had a range of 0 to 1.00. The more decisions that were made by the subordinates, the higher the value of the index and the greater the principal's propensity to delegate decision making.

Dependent Variables

During a regular faculty meeting, one-quarter of the faculty, selected at random, responded to measures of the dependent variables—hierarchy of authority, job satisfaction, and loyalty to the principal—while the rest of the faculty completed other research measures not used in this study. Virtually all teachers selected completed the research instruments. A total of 922 (96 percent) of the 965 forms completed was usable. Individual scores were combined into a mean score for each school; hence, the school scores were used in all hypothesis testing.

Hierarchy of Authority was measured with a fourteen-item index developed from the work of Hall (1963), McKay (1964), and Aiken and Hage (1966). Teachers were asked to respond to each item along a five-point response scale of "always," "often," "occasionally," "seldom," or "never" (always=5). Examples of items were: "Any decision I make has to have my superior's approval," "I have to ask the principal before I do almost anything," and "There can be little action until a supervisor approves a decision." The reliability for the measure was supported by an alpha coefficient of .89 for the sample. Construct validity of the measure has been demonstrated in a number of studies (Sousa & Hoy, 1981; Punch, 1969).

Job Satisfaction was operationalized using an eight-item index developed by Aiken and Hage (1966). Teachers were asked to answer the questions in terms of their degree of satisfaction along a four-point scale from 4 (very satisfied) to 1 (very

dissatisfied). Samples of questions included the following: “How satisfied are you that you have been given enough authority by administrators to do your job well?” “How satisfied are you with your present job when you compare it to similar positions?” and “How satisfied are you with your present job in light of career expectations?” Alpha coefficients for the index have been consistently about .80 (Hoy, Blazovsky, Newland, 1983); in the present study the alpha was .83. Likewise construct validity of the measure has been supported in a number of studies (Aiken & Hage, 1966; Hoy, Blazovsky, & Newland, 1983).

Loyalty to the Principal was measured with an eight-item index adapted by Hoy and Williams (1971) from the work of Murray and Corenblum (1966). The following questions are examples of the items to which teachers responded: “If you had a chance to teach for the same pay in the school under the direction of another principal, how would you feel about moving?” “Generally speaking, how much confidence and trust do you have in your principal?” “All in all, how satisfied are you with your principal?” Construct validity for the measure has been supported in several studies (Hoy, Tarter, & Forsyth, 1978; Hoy & Rees, 1974; Hoy & Williams, 1971). Reliability has been consistently high with alpha coefficients in the .90 range; in this study the alpha was .92.

Results

Although 19 principals in the present sample delegated some decisions to subordinates, 36 principals delegated no decision making authority to teachers in the ten important decision making areas. Thus, the independent variable was dichotomized into delegation or no delegation of decision making, and the three hypotheses guiding the study were tested using analysis of variance procedures.

TABLE 1
SUMMARY AND ANALYSIS OF VARIANCE DATA: THE RELATIONSHIP OF
HIERARCHY OF AUTHORITY AND PRINCIPALS' DELEGATED DECISION-MAKING

	Delegation		No Delegation		
	19		36		
Number					
Mean	2.84		3.20		
Standard Deviation	.31		.25		
Source	Sum of Squares	df	Mean Square	F	P
Between Groups	1.621	1	1.621	18.9	<.01
Within Groups	4.553	53	0.086		
Total	6.174	54	0.114		

All three hypotheses were supported. Principals who delegated their decision-making authority had teachers who described the authority structure of the school as significantly less rigid than teachers who had principals who did not delegate ($F=18.9, p<.01$). Likewise, principals with a propensity to delegate had teachers

who were significantly more satisfied with their job ($F=10.4, p<.01$) and who were significantly more loyal ($F=5.6, p<.05$) than principals who did not delegate. The analysis of variance data are summarized in Tables 1-3.

TABLE 2
SUMMARY AND ANALYSIS OF VARIANCE DATA: THE RELATIONSHIP OF
JOB SATISFACTION AND PRINCIPALS' DELEGATED DECISION MAKING

		Delegation		No Delegation	
Number		19		36	
Mean		2.93		2.71	
Standard Deviation		.23		.25	
Source	Sum of Squares	df	Mean Square	F	P
Between Groups	0.606	1	1.606	10.4	<.01
Within Groups	3.088	53	0.058		
Total	3.694	54	0.069		

TABLE 3
SUMMARY AND ANALYSIS OF VARIANCE DATA: RELATIONSHIP OF TEACHER
LOYALTY AND PRINCIPALS' DELEGATED DECISION MAKING

		Delegation		No Delegation	
Number		19		36	
Mean		3.57		3.28	
Standard Deviation		.48		.43	
Source	Sum of Squares	df	Mean Square	F	P
Between Groups	1.100	1	1.100	5.6	<.05
Within Groups	10.460	53	0.197		
Total	11.560	54	0.214		

Summary and Discussion

The findings in this study suggest that the willingness of principals to delegate important decisions to teachers (and other subordinates) can result in potential benefits for the school organization. First, teachers are less likely to perceive a

strong authority structure and thus may be more apt to identify with the goals and objectives of the school. Second, teachers gain a greater sense of job satisfaction, which tends to produce improved attitudes toward their work and the people with whom they work. Third, teachers are more inclined to exhibit loyalty to their principals.

The confirmation of all three hypotheses supports the validity of the index of delegation of decision making. The measure of propensity to delegate was determined by first establishing the important and specific internal decisions that the principal had the authority to make and then ascertaining which, if any, of those decisions were given to subordinates. The assumption was made that principals could delegate legitimately only decisions for which they had authority. Moreover, if principals shared important decisions, it was expected that teachers would view the hierarchy of authority as more flexible, an assumption that was supported by the findings. Further, we postulated that if teachers were granted authority to make important decisions, then they would feel a greater sense of self-attainment and influence thereby increasing their job satisfaction. Again, the findings supported the underlying theory.

A basic challenge facing all principals is to find ways to extend the scope of their influence over teachers beyond the narrow limits of formal authority flowing from their position. Barnard (1938, p. 170) has cogently suggested that only when the authority of leadership is combined with the authority of position will superiors be able to induce subordinates to comply with orders outside their bureaucratic zone of indifference. The informal organization is an important source of leadership, and principals who can command loyalty from teachers seem to have a distinct advantage in enlarging their authority over teachers (Hoy & Rees, 1974). The findings of this study suggest that principals who have the propensity to delegate decisions have a significant edge in commanding teacher loyalty. When principals entrust to teachers the authority to make important decisions, it seems likely teachers reciprocate by developing norms of allegiance, trust, and loyalty to the principal. Moreover, commanding subordinate loyalty has been linked with productivity. Blau and Scott (1962, p. 144) report that, "workers' productivity was largely associated with factors that were related to loyalty perhaps because supervisors who had won the loyal support of their subordinates were most successful in commanding willing compliance with their directives and in stimulating effort in their work group." Yet the results of this study show the principals are extremely reluctant to share their authority, with 36 of the 55 principals not delegating in the ten important areas.

The evidence, however, continues to grow that secondary school principals have significant authority to run their schools. For example, consistent with the findings of this study, two national studies of secondary school principals (Abramowitz & Tenenbaum, 1978; Brynes et al., 1978) show that the majority of principals select their own teachers, with only confirmation from the central office. Furthermore, most principals have major authority in nearly all matters that deal with the school's internal operation. Despite this extensive power of many principals, and despite the apparent benefits derived from entrusting some authority to subordinates, most principals in this sample (65 percent) had *no* propensity to delegate decisions to teachers in areas which teachers seemed to possess both expertise and a personal stake in the outcome. Perhaps principals consulted with teachers or jointly involved them in the process of making decisions, but most did not assign even such tasks as deciding to introduce a new course or program, or deciding

which suppliers or materials to use—let alone delegate decisions involving the selection of teachers or the internal allocation of funds.

Perhaps many principals typically are more concerned with their own position, status, and career than they are with the professional growth and development of subordinates. Perhaps some principals are too insecure to share power. Thirty-five percent of the principals in this study, however, did transfer significant decision making authority to their subordinates. Apparently, some principals do see the benefits. It could be hypothesized that principals who do delegate major decision making authority are more secure and more likely to view their subordinates as professionals than those who do not. Of course, the test of such hypotheses awaits further empirical investigation.

Clearly, some principals who did not delegate allowed subordinates to participate in joint decision making, but they reserved the final decisions for themselves. This fact points to a practical distinction between delegated and consultative decision making. Although many principals are willing to inform and to solicit information from subordinates on pending decisions, they do not delegate the decision. The result is that teachers begin to doubt the benefits of such participation and voice skepticism over whether their so-called involvement makes any difference (Duke et al., 1980). Apparently, teacher participation in decision making that lacks influence may be less beneficial than no teacher involvement at all.

This study also points to important theoretical distinctions among delegated, consultative, and joint decision making. Delegation of decision making maximizes participation; consultation minimizes participation. The term participation is simply too ambiguous and broad. The time has come to more clearly specify the meaning of involvement in decision making. The concept of delegation is a step in that direction. Moreover, the index developed in this research should provide researchers with a useful tool for the continued study of a concept that has both theoretical and practical significance.

Notes

1. An earlier version of this paper was presented at AERA in Montreal (1983).
2. The Aston approach was first modified for use in public schools by a group of researchers from the University of Alberta (Newberry, 1971; Kelsey, 1973; Holdaway, Newberry, Hickson, & Heron, 1975; Sackney, 1976).

References

- Abramowitz, S., & Tenenbaum, E. (1978). *High school '77*. Washington: National Institute of Education.
- Aiken, M., & Hage, J. (1966). Organizational alienation: A comparative analysis. *American Sociological Review*, 31, 497-507.
- Barnard, C. (1938). *The functions of the executive*. Cambridge, MA: Harvard University Press.
- Blau, P. M., & Scott, W. R. (1962). *Formal organizations*. New York: Chandler.
- Brynes, D. R. et al. (1978). *The senior high school principalship: The national survey (Vol. 1)*. Reston, VA: National Association of Secondary School Principals.
- Conway, J. A. (1976). Test of linearity between teachers' participation in decision making and their perceptions of their schools or organizations. *Administrative Science Quarterly*, 21, 130-139.

- Crockenberg, V., & Clark, W. W. (1979). Teachers' participation in school decision making: The San Jose teacher involvement project. *Phi Delta Kappan*, 61, 115-118.
- Cross, R. (1980). A description of decision-making patterns of school principals. *Journal of Educational Research*, 73, 154-159.
- Dachler, H. P., & Wilpert, B. (1978). Conceptual dimensions and boundaries of participation in organizations: A critical evaluation. *Administrative Science Quarterly*, 23, 1-39.
- Davis, K. (1963). The case for participative management. *Business Horizons*, 7, 45-50.
- Davis, K. (1967). *Human relations at work: The dynamics of organizational behavior*. New York: McGraw-Hill.
- Duke, D. L., Showers, B. K., & Imber, M. (1980). Teachers and shared decision making: The costs and benefits of involvement. *Educational Administration Quarterly*, 16, 93-106.
- Hall, R. H. (1963). The concept of bureaucracy: An empirical assessment. *American Journal of Sociology*, 63, 32-40.
- Healey, R. W. (1980). Principals' perceptions of faculty involvement in decision making in public elementary schools of Illinois. *Illinois School Research and Development*, 16, 91-97.
- Holdaway, E. A., Newberry, J. F., Hickson, D. J., & Heron, R. P. (1975). Dimensions of organizations in complex societies: The educational sector. *Administrative Science Quarterly*, 20, 37-59.
- Hoy, W. K., Blazovsky, R., & Newland, W. (1983). Bureaucracy and alienation: A comparative analysis. *Journal of Educational Administration*, 21, 109-210.
- Hoy, W. K., Newland, W., & Blazovsky, R. (1977). Subordinate loyalty to superior, esprit, and aspects of bureaucratic structure. *Educational Administration Quarterly*, 13, 71-85.
- Hoy, W. K., & Rees, R. (1974). Subordinate loyalty to immediate superior: A neglected concept in the study of educational administration. *Sociology of Education*, 47, 268-286.
- Hoy, W. K., Tarter, C. J., & Forsyth, P. (1978). Administrative behavior and subordinate loyalty: An empirical assessment. *Journal of Educational Administration*, 16, 29-38.
- Hoy, W. K., & Williams, L. B. (1971). Loyalty to immediate superior at alternate levels in public schools. *Educational Administration Quarterly*, 7, 1-21.
- Isherwood, G. B., & Taylor, R. M. (1978). Participatory decision making via school councils. *The High School Journal*, 255-270.
- Kelsey, J. G. T. (1973). *Conceptualization and instrumentation for the comparative study of secondary school structure and operation*. Doctoral dissertation, University of Alberta, Edmonton.
- Kunz, D. W., & Hoy, W. K. (1976). Leadership style of principals and the professional zone of acceptance of teachers. *Educational Administration Quarterly*, 12, 49-64.
- Leavitt, H. J. (1965). Applied organizational change in industry: Structural, technological and humanistic approaches. In J. G. March (Ed.), *Handbook of Organizations*. Chicago: Rand McNally.
- Likert, R. (1967). *The human organization: Its management and value*. New York: McGraw Hill.
- Locke, E. A., & Schweiger, D. M. (1979). Participation in decision making: One more look. *Research in Organizational Behavior*, 1, 265-339.
- Lowin, A. (1968). Participative decision making: A model, literature critique, and prescription for research. *Organizational Behavior and Human Performance*, 3, 68-106.
- McKay, D. A. (1964). An empirical study of bureaucratic dimensions and their relationship to other characteristics of school organizations. Unpublished doctoral dissertation, University of Alberta, Edmonton, Alberta.
- Murray, V. V., & Corenblum, A. F. (1966). Loyalty to immediate superior at alternate hierarchical levels in a bureaucracy. *American Journal of Sociology*, 22, 77-85.

- Newberry, J. F. (1971). *A comparative analysis of the organizational structures of selected post-secondary educational institutions*. Unpublished doctoral dissertation, University of Alberta, Edmonton.
- Pugh, D. S., & Hickson, D. J. (1978). *Organizational structure in its context*. Westmead, Farnborough, Hants, England: Saxon.
- Punch, K. F. (1969). Bureaucratic structure in schools: Towards redefinition and measurement. *Educational Administration Quarterly*, 5, 43-57.
- Sackney, L. E. (1971). *The relationship between organizational structure and behavior in secondary schools*. Unpublished doctoral dissertation, University of Alberta, Edmonton.
- Sashkin, M. (1976). Changing toward participative management approaches: A model and methods. *Academy of Management Review*, 1, 75-86.
- Schacht, R. (1970). *Alienation*. Garden City, New York: Doubleday.
- Schregle, J. (1970). Forms of participation in management. *Industrial Relations*, 9, 112-117.
- Sousa, D. A., & Hoy, W. K. (1981). Bureaucratic structure in schools: A refinement and synthesis in measurement. *Educational Administration Quarterly*, 17, 21-39.
- Strauss, G., & Rosenstein, E. (1970). Workers' participation: A critical view. *Industrial Relations*, 9, 197-214.
- Tannenbaum, A. S. (1974). Systems of formal participation. In G. Strauss et al. (Eds.), *Organizational behavior: Research and issues*. Madison, WI: Industrial Relations Research Association.

PERSPECTIVES

The Theory-Practice Relationship in Teacher Education: Need There Be Any?

Peter J. Miller
and
Gerald D. Taylor
University of Alberta

Perhaps the most persistent and vexing problem area in teacher education programs is the place occupied by the so-called theoretical studies in education. No one seriously questions whether teacher trainees ought to develop appropriate and adequate subject matter expertise through their university studies. Similarly, no one argues against the need for teacher trainees to be given extensive opportunities to practice teach before they are certified as competent teachers. Subject matter expertise and practical experience—these are the obvious and undisputed elements of teacher education programs. But about the place and purpose assigned to professional theoretical course work and the relationship that this abstract theoretical study of education ought to have to the actual performance of the teacher in the classroom, there exists no such agreement.

In recent years this professional theoretical component of teacher preparation programs has been the subject of considerable criticism and dissatisfaction. Politicians, administrators and students have increasingly demanded an initial preparation which fosters not so much the theoretical abstract understanding of education but rather the specific skills and competencies which beginning teachers are believed to need if they are to perform effectively in classrooms as they presently exist. In response to this demand, institutions of teacher education have

Peter Miller (PhD, Alberta) is Dean of Student Services at the University of Alberta and a Professor in the Department of Educational Foundations. His major teaching and research area is the history of childhood as this relates to the contemporary belief systems of parents and teachers.

Gerald Taylor (PhD, Toronto) is an Associate Professor at the University of Alberta in the Department of Educational Foundations. His major research and teaching interests include the sociology of learning, the sociology of teaching, and school-community relations.

tended to center their programs on the most directly practical aspects—instructional and management skills and various competencies developed within an extended practicum framework. With this emphasis on teaching competencies and the practicum, both geared towards entry level certification, the more abstract theoretical courses in teacher education programs have faced considerable pressure, either to show a payoff that is directly related to classroom teaching or to step aside in favor of more skill and practice oriented experiences. Even highly theoretical education courses, which on the surface appear to have little connection with the teaching act (e.g., history, philosophy and sociology of education), have been evaluated in terms of their perceived relevance to teaching performance and classroom effectiveness. Within this general view of teacher education, then, it is held that the relationship between theoretical course work and the instructional and management skills demonstrated in the practicum experience must be a direct one. In effect, university based theoretical course work is expected to affect—that is, to be causally related to—performance in the classroom, both as a student and as a certified teacher. And when it is shown or believed that this has not occurred, the typical response is to seriously question the relevance of such course work and hence its value to initial teacher preparation.

What research has been reported, however, suggests that there is little relationship between performance in theoretical course work and performance in teaching. Jencks (1972, p. 187) conclude from a number of studies relating professionals' performance on the job to college grades that this relationship is much weaker than commonly supposed. In the case of elementary and secondary school teachers, they point out that college grades and observer ratings of teacher performance correlate only between 0.2 and 0.3. Other researchers have uncovered similar statistical trends within the field of teacher education. Emanuel, Larimore, and Sagan (1975), in a study of 42 secondary social studies majors, found that student teaching performance was not significantly correlated with final grades in any of seven different education courses. The correlation coefficients reported in their study ranged from a low of 0.05 between grades in a general teaching methods course and student teaching performance to a high of only 0.28 between the latter and grades in a senior social studies methods course. And although Twa and Greene (1980) recently found, in a study of 139 teacher trainees at the University of Lethbridge, that students' grade point average in their teaching major was significantly correlated with their performance in practice teaching, the magnitude of this correlation was only 0.26 (p. 8). Similarly, while Emerson, Elford and Scaldwell's (1982) study of 112 Althouse teacher trainees found a positive association between practicum success and performance in foundational studies (i.e., theoretical course work in psychology, philosophy and sociology), the phi coefficients ranged from a low of 0.03 in the case of philosophy to a high of only 0.34 in psychology. Further evidence in support of these findings is provided by a recent study at the University of Alberta conducted by Taylor and Miller (1984) in which little correlation was found to exist between grades achieved in professional courses during a professional year and practice teaching ratings.

These findings are in line with those of other researchers, findings which have led Burton to conclude from the literature that performance in teacher training courses bears little relationship to success in teaching.¹ Indeed, the evidence currently available tends to support the observation drawn by Cortis and Satterly (1968) who, in reference to teacher education in England, concluded that there is an apparent dichotomy between the criteria for success within the academic/theoreti-

cal component of teacher education and the criteria needed for success within the practicum component. Good students apparently do not necessarily make good teachers. Understandably, such evidence tends to increase the concern expressed by those who have grown impatient with theory and skeptical of its value to practice; it also seriously weakens the position of those who hold tenaciously but precariously to the view that theory must be the basis on which professionals, particularly teachers, build their practice.

For many people working in the field of education this weak association between performance in theoretical professional courses and performance in teaching will come as no surprise. Students who have completed at least some practice teaching, classroom teachers, and school personnel engaged in the hiring of teachers frequently comment that the qualities, aptitudes and skills required for successful performance as a university student and successful performance as a classroom teacher are sufficiently different that one ought not to expect much of a relationship between the two. However, for those concerned with the design of teacher education programs, this weak relationship is disturbing and of crucial importance since it raises questions about the appropriateness and, indeed, the validity of the current approach to teacher education. It casts serious doubt, in fact, on the assumption that existing structures, mechanisms, and strategies in teacher education can, under even the very best conditions, ever effectively integrate theory and practice.

In itself the well documented lack of any relationship between performance in theory and performance in practice carries with it no implications for teacher education. It is merely an observed and, for most people not so surprising, phenomenon. When, however, this phenomenon is considered in the light of a particular model of teacher education, with all its various value commitments and assumptions about the priorities of professional education, then a number of important conclusions can be drawn and, more importantly, recommendations formulated. Consider in this connection the significance of this observed lack of a theory-practice performance relationship in the light of three models of teacher preparation.

A weak association between the two modes or domains of performance is unsettling to all those who believe that university-based initial teacher education, whether it culminates in a degree or not, should focus primarily on the development of the entry level teaching skills needed for effective performance in a conventional classroom. But it is most disturbing to those who insist that it is the teaching practice or practicum experience that is of fundamental importance in the development and firming up of these necessary teaching competencies. For this practice-driven or competency-based model, the practicum component of initial teacher education is, of course, central to the entire program. The Bachelor of Education degree program's professional course work thus tends to be conceived, designed and evaluated in light of the skills and competencies needed by teacher trainees to successfully complete practice teaching and thus become eligible for initial certification as beginning classroom teachers.

For those committed to this practice-driven model of a Bachelor of Education degree program, when little apparent association exists between being a good student and being a good teacher, one of two related conclusions can be drawn. First, the theoretical courses may be poorly designed or out of touch with the realities of classroom teaching, and hence stand in need of basic change in content or mode of delivery. More fundamentally, it may be concluded that conventional formal professional course work has lost its validity or *raison d'être*. The implications of such a position are, as Emanuel, Larimore and Sagan (1975) point out, that

Colleges of Education would be well advised to examine closely their teacher training programs and decide whether all of these courses are worth keeping in the program or whether their content should be integrated with actual classroom experiences in order to establish a greater relationship between theory and practice. (p. 248)

From the standpoint of this model, then, the theory component must be more closely molded around the practicum since the latter operates within the real world of the classroom. And this can be accomplished most effectively, Emanuel, Larimore, and Sagan (1975) inform us, by replacing conventional theory courses with what they refer to as "experientially-based" courses (p. 248). These, they argue, will integrate the objectives of the courses contained in teacher preparation programs, demonstrate to the trainees the utility of theory to practice, and strengthen the relationship between course performance and performance in student teaching. These benefits will result because under such conditions the two behaviors being compared become very similar.

Advocates of this model hold, therefore, that the theory component must be re-fashioned so that the competencies needed and the criteria of evaluation in theoretical studies converge with those of the practicum. Reform of teacher education is envisaged, in effect, primarily in terms of redesigning the theoretical elements since it is their relevance, utility and hence legitimacy that are in question. Conversely, from this perspective, improvements of the practice component entail, for the most part, an extension of the practicum rather than any alteration of its present nature and purpose.

But to be committed to an initial certification model of teacher education and thus of the Bachelor of Education program, it is not also necessary to be committed to a practicum-driven or competency-based model. Consider, for example, a model of teacher preparation which runs counter to it and as such leads to substantially different conclusions about the weak association between being a good student and being a good teacher. This alternative perspective, which may be termed the theory-driven model, holds that the practicum experience is not the most important element of the teacher education program; it should not, therefore, dictate the nature and purposes of theoretical studies during the initial preparation of teacher candidates. On the contrary, exponents of this perspective argue that sound practice relies heavily on a strong foundation not only of general knowledge derived from liberal arts but also on in-depth understandings of teaching and learning that can be fostered by professional education courses. Hence, teacher trainees must engage extensively in the study of "the most recent and well-founded knowledge in the disciplines of educational study" (Taylor, 1981, p. 3). They must also reflect on the nature of this knowledge and its implications for learning and teaching. And no less significantly, they must demonstrate an awareness, understanding and appreciation of these theoretical underpinnings of sound teaching practice. The importance of such understanding cannot be overstated since it is this body of educational theory that provides the basis on which to build, in Joyce's (1981) terms, "the intellectual and interpersonal tools to continue to study and expand the range of highly developed skills" (p. 2) required of teacher candidates if they are to be permanently committed, as Taylor (1981) puts it, "to personal learning and professional development" (p. 3).

From the perspective of this theory-driven model of initial preparation, the practicum is still of central importance; its primary purpose, however, is to serve as a bridge between educational theory and practice. As expressed in the Government of Alberta's Theory to Practice report (1981), practicum experience should be de-

signed to “bring theoretical and practical aspects of undergraduate education into a clear and interacting type of relationship and to aid students in applying what they learn theoretically to simulated and actual situations” (p. 14). In effect, the practicum is envisaged as a kind of laboratory experience in which the theory studied in course work can be, in some sense, tested in the classroom.

But if the practicum is to help foster among teacher candidates an appreciation of the relevance and utility of theory to actual classroom teaching and, moreover, culminate in an ability to translate theoretical concepts into teaching practice, it must be designed specifically with these objectives in mind. In particular, to again borrow from the Government of Alberta’s Theory to Practice report (1981), the practicum must provide ample opportunity for student trainees to “create, innovate, test and revise theories and methodologies” (p. 15). Moreover, the analysis of practice within it must be, according to Taylor (1981, p. 10), “systematically conducted in terms of concepts and theories which feature in other areas of the professional sequence” comprising the teacher education program.

However, according to proponents of this theory-driven model, it is highly doubtful that any clear articulation between theory and practice can be achieved within a program of teacher education which is firmly tied to the certification of entry level teaching competencies and skills—that is, when it is practice-driven. According to Joyce (1981, pp. 1-2), initial preparation should prepare teacher candidates:

1. For the lifelong study of the world, the self and academic knowledge.
2. The lifelong study of teaching.
3. To participate in school renewal efforts including the creation and implementation of innovations.
4. To approach the generic problems of the workplace (the school and the classroom).

Such objectives are extremely difficult to achieve, however, when conventional practicum experiences have as their goal socializing teacher candidates to the role of teachers in conventional classrooms. For as Joyce notes, the student teaching experience tends to introduce and initiate prospective teachers into the conventional wisdom, procedures and values of the current approach to practice rather than systematically to provide for the opportunity to test out and implement emerging or more progressive strategies suggested in theoretical course work. Moreover, during the practicum many teacher candidates also develop an attitude of pragmatism which turns them “from university and gradually alienates them from the intellectual roots of their profession” (p. 12). They tend, therefore, to “become concerned with quick and easy solutions that will make immediate differences to their classroom rather than the lifelong study of the more complex and interesting ways of approaching children” (p. 12); and they become “more constrained in their behavior, more authoritarian and more resistant to change” (p. 9). Operating under the pressure of summative credential-oriented evaluation, they come to “feel tremendous stress . . . [which] has the effect of driving the[ir] self-concept . . . toward (in Maslow’s terms) a survival orientation” (p. 9).

For proponents of this theory-driven model, therefore, if the abilities, skills, and competencies required for success in theoretical studies are dramatically different from those needed for successful performance in the conventional practice, what stands in need of serious revision are the latter, not the former. In other words, the evidence indicating that good students do not necessarily make good teachers is in-

terpreted to mean not that theoretical course work is irrelevant to teaching effectiveness in the practicum experience, but rather that the practicum, and indeed, school practice itself, are being conducted and evaluated inappropriately.

From the lack of any strong relationship between performance in theoretical course work and performance in practice can be drawn, then, two widely differing sets of implications for teacher education, the one resulting in a critique and proposed revision of the theory element, the other in a critique and proposed revision of contemporary practicums, and indeed, of classroom practices. Both the practicum-driven and the theory-driven models, however, are similar in that they are both based upon a conviction that theory and practice should be both theoretically and empirically related. Both insist, therefore, that entry level certification and the Bachelor of Education degree should be granted upon successful completion of a program in which professional theoretical course work has been, in some fashion, causally related to the practicum or student teaching experience. In this respect both models are based on convictions and assumptions about professional education which, according to the evidence at least, have little basis in reality. Nor is it clear if a knowledge of theory, that is, of the various concepts and ideas dealt with in professional course work, can be even logically related to the various competencies, skills and attitudes which are required for successful performance in classrooms as they presently exist.

There are thus good grounds for suggesting that faculties of education should consider an alternative model of professional education, one that stresses not the logical and/or empirical interdependence and interrelationship of theoretical studies and certification-oriented practical training but rather their discreteness or separateness. Each may well be of fundamental importance in professional training but this does not mean that they are therefore causally related. Indeed, each should be seen as having its own rationale and each its own unique and appropriate criteria for evaluating success or failure.

It is important in this connection to separate out what can be achieved during a university degree program from what can be achieved in an apprenticeship or internship setting. What a Bachelor of Education degree program provides is knowledge. To begin with, and most importantly, it provides a knowledge of the subject matter to be taught; and about the value of this kind of knowledge there is no debate. But it also typically includes elements of theoretical professional studies which provide the basis of an informed understanding of the professional activity in all its contexts. The professional payoff of such knowledge is considerable. To begin with, in the western world this kind of understanding is generally held to be the distinguishing characteristic of a professional activity as opposed to a craft or trade. It is thus closely tied to both income level and occupational status. It is also an indispensable prerequisite for advanced study and research in a professional field and the basis of lifelong learning to which teachers should be committed. Finally, it is directly related to an ability to critically evaluate professional practice, not only within the narrow terms of professional concerns but also within the larger social context; it is thus linked to a willingness to experiment and innovate. It should be noted that this kind of knowledge and the attitudes and values it embodies exist quite independently of the competencies needed for success in a conventional classroom.

The criteria for demonstrating such knowledge in a university setting are well known and need no further elaboration. They are certainly clearly different from those which govern the evaluation of performance in a practicum activity which has

as its primary goals the development of needed job competencies and, equally important, socialization of the trainee into the professional or conventional order of things. To allow these two discrete areas of professional education to spill over into each other and to exercise a strong influence on each other is to compromise the integrity of both. On the one hand, it seriously weakens the quality of what is essentially an academic activity while, on the other, it intellectualizes unnecessarily what is basically a practical and socializing experience.

The implications of such conclusions for the professional education of teachers are profound. The Bachelor of Education degree should be freed from its current preoccupation with entry level skills and become an academic credential awarded on the basis of academic considerations. The theoretical or professional studies component of the degree program would, of course, involve students in a variety of experiences that would take place in school settings. These laboratory experiences, however, would not be viewed as preparation or training for effective functioning in a conventional classroom—that is, as a method or means of acquiring a particular repertoire of specific teaching competencies. Rather, the nature and function of these experiences would be determined by the academic requirements of theoretical course work.

Whether school boards and teachers would be willing to throw open their classrooms to the experimental activities and inevitably critical gaze of university instructors and their students is doubtful. And, of course, it is not clear whether this kind of use of conventional classrooms could be justified or, indeed, whether it would be worthwhile. It may be that faculties of education should create their own laboratory schools in which students could observe a variety of classroom phenomena, critically evaluate the appropriateness of particular kinds of procedures, and experiment in all areas of teaching without any concern that their efforts will be judged on grounds that have to do with conventional norms of teaching effectiveness.

The Bachelor of Education degree, when conferred, would certify that the holder had achieved a certain level of general knowledge, subject matter expertise and a theoretical understanding of the various aspects of education. It would thus be similar in kind to the degrees awarded in other professional fields, notably business, law and medicine. Its possession would permit the holder to apply for a further program of professional training, an internship, to be conducted in a school setting. When successfully completed, the internship would lead to certification as a professional teacher. The details of such internship programs need not concern us here, although they would probably have a good deal in common with existing conventional extended practicum programs. They would have as their goal the training of the intern in all the needed job competencies required for effective functioning in the contemporary classroom. The internship training program could be developed on the assumption that the interns were familiar with the theoretical aspects of education. It would not, however, need to be designed in such a way that its relevance or practicality would be compromised by importing into it any abstract theoretical distractions.

Whether this separation of theory and practice—of academic and non-academic elements—in teacher education will produce better teachers, we have no way of knowing. But it seems likely that such programs will produce teachers who will be better educated during their degree program, and better trained during their internship period. And these are surely steps in the right direction.

Notes

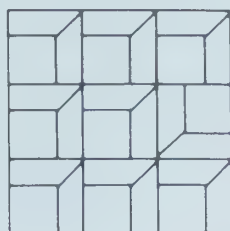
1. Cited in Emanuel, Larimore and Sagan (1975, p. 246).

References

- Cortis, G. A., & Satterly, D. J. (1968). Personality profiles of student teachers. *British Journal of Psychology*, 32(2), 115-122.
- Emanuel, J., Larimore, D., & Sagan, E. (1975). The relation of selected academic program variables to student teaching performance. *Contemporary Education*, 46(4), 245-248.
- Emerson, G., Elford, G., & Scaldwell, W. (1982). Who succeeds in practice teaching. *News Quarterly*, 1(5), 1-5.
- Faculty of Education. (1983). *Handbook for practicum programs in the elementary schools, 1983-84*. University of Alberta.
- Government of Alberta. (1981). *Theory to practice: Report of the committee to evaluate the extended practicum at Alberta universities*.
- Jencks, C. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Harper & Row.
- Joyce, B. R. (1981). The phoenix plan: Essential reform in teacher education. In Government of Alberta, *Theory to practice: Report of the committee to evaluate the extended practicum at Alberta universities*. Appendix B.
- Taylor, G. D., & Miller, P. J. (in press). Professional course work and the practicum: Do good students make good teachers? *Canadian Journal of Education*.
- Taylor, W. (1981). Paper for the committee on the evaluation of teacher education programs in Alberta. In Government of Alberta, *Theory to practice: Report of the committee to evaluate the extended practicum at Alberta universities*. Appendix A.
- Twa, R. J., & Greene, M. (1980). Prediction of success in student teaching as a criterion for selection in teacher education programs. *The Alberta Journal of Educational Research*, 26(1), 1-13.

Table of Contents, Volume XXX, 1984

- Allison, D. E. *The effect of homogeneous vs. heterogeneous matching-item format on test performance and reliability.* No. 2, June, 91-94.
- Belanger, J., and Martin, R. G. *The influence of improved reading skill on writing skill.* No. 3, September, 194-212.
- Carey, S. T., and Cummins, J. *Communication skills in immersion programs.* No. 4, December, 270-283.
- Dash, U., and Maguire, T. *A detailed analysis of group differences on the California short-form test of mental maturity between 1956 and 1977.* No. 2, June, 95-103.
- Evans, P., Mickelson, N., and Smith, N. C. *Computer-assisted instruction in college composition.* No. 4, December, 311-319.
- Friesen, D., Holdaway, E. A., and Rice, A. W. *Factors contributing to the job satisfaction of school principals.* No. 3, September, 157-170.
- Gayle, G. M. H. *Second language teaching styles: Making or breaking attitudes.* No. 1, March, 49-67.
- Gunderson, L. *One last word list.* No. 4, December, 259-269.
- Houle, R., and Montmarquette, C. *An empirical analysis of loans by school libraries.* No. 2, June, 104-114.
- Hoy, W. K., and Sousa, D. A. *Delegation: The neglected aspect of participation in decision making.* No. 4, December, 320-331.
- Levin, B. *Development and initial results of a longitudinal secondary follow-up study.* No. 1, March, 1-13.
- Maynes, F. J., and Ross, J. A. *Integrating school curricula with social concerns.* No. 3, September, 179-193.
- Merritt, F. M., and McCallum, R. S. *The relationship between simultaneous-successive processing and academic achievement.* No. 2, June, 126-132.
- Miller, P. J., and Taylor, G. D. *The theory-practice relationship in teacher education: Need there be any? (Perspectives)* No. 4, December, 332-339.
- Mueller, H. H., Dash, U. N., Matheson, D. W., and Short, R. H. *WISC-R subtest patterning of below average, average, and above average IQ children: A meta-analysis.* No. 1, March, 68-85.
- Mueller, H. H., Mancini, G. J., and Short, R. H. *An evaluation of the diagnostic efficiency of the Wechsler Intelligence Scale for Children—Revised.* No. 4, December, 299-310.
- Nixon, M. T., and Bumbarger, C. S. *Prospective teachers: Views of programs and needs.* No. 3, September, 226-237.
- Pansegrau, M. *Teachers' perspectives on inservice education.* No. 4, December, 239-258.
- Rampaul, W. E., Singh, M., and Didyk, J. *The relationship between academic achievement, self-concept, creativity, and teacher expectations among Native children in a northern Manitoba school.* No. 3, September, 213-225.
- Robinson, N., and Stacey, C. *Serving on the school board: A political apprenticeship for higher office.* No. 2, June, 115-125.
- Staab, C. F. *Student teachers' perceptions of their competence versus the perceptions of university supervisors and sponsor teachers.* No. 4, December, 284-298.
- Walberg, H. J., Singh, R. P., and Tsai, S. L. *The achievement productivity of psychological variables: A replication and extension in Rajasthan, India.* No. 3, September, 171-178.
- Westbury, M. *The link between universities and employment in Alberta.* No. 1, March, 38-48.
- Whyte, L. A. *Characteristics of learning disabilities persisting into adolescence.* No. 1, March, 14-25.
- Wideen, M. F., and Hopkins, D. *Supervising student teachers: A means of professional renewal?* No. 1, March, 26-37.



Journal of Education

Published by
Boston University
School of Education

Winter 1983 Volume 165, Number 1

Literacy and Ideology

focusing on the theoretical relationship of reading, literacy, and ideology with specific studies of the French immersion program in Canada and the literacy programs in the Cape Verde Islands and Guinea-Bissau. Papers by

Paulo Freire
Henry Giroux
James Donald
Paul Olson
Donaldo Macedo

Spring 1983 Volume 165, Number 2

The Dialectics of Evaluation

focuses on the social and political goals of educational environments and their implications in evaluation and testing: Studies of the pre-school environment, student classification, and testing considered in the light of gender and social policy.

Winter 1984 Volume 166 Number 1

Class, Gender, and Schooling

Papers by

Jean Anyon
Madeline Arnot
Miriam David
Rosemary Deem
Jane Gaskell
Ann Manicom

Requests for single copies or subscriptions to:

Journal of Education
Boston University
605 Commonwealth Avenue
Boston, Massachusetts 02215

Single issue:	\$ 8.00
Individual one year subscription:	\$16.00
Institutional one year subscription:	\$21.00

PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in triplicate and are not to exceed 25 pages in length.
2. An abstract of approximately 100 words in length, typed on a separate page, is to be included.
3. The author's name and affiliation should appear only on a separate title page to ensure anonymity in the reviewing process; the title of the paper should appear on the first page of the manuscript as a means of identification.
4. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publication Manual.
5. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
6. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
7. Sources should appear in parentheses following the reference, citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References*.
8. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 7 above.
9. Spelling shall conform to *Webster's New Collegiate Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
10. In matters of style, the third edition of the APA Publication Manual is considered definitive.

